microelectronics group

Lucent Technologies
Bell Labs Innovations

TECO3264 32-Channel Echo Canceller Information Manual

Introduction

This Information Manual supplements the detailed technical information in the August 1999 *TECO3264 32-Channel Echo Canceller* Data Sheet (DS99-241PDH) (which contains parameter values, microprocessor interface, electrical characteristics, etc.) with a brief description of the TECO3264 echo canceller integrated circuit, discussions of network electrical echo canceller applications, and of background information on network characteristics that affect echo cancellers. Both traditional long-haul telephone networks and newer voice telephony transport systems are covered. The third document in this series is the TECO3264 Evaluation Board System User Manual which describes a stand-alone test system for the TECO3264.

The TECO3264 echo canceller has been developed to provide an efficient, economical, state-of-the-art solution to electrical echoes in the telephone network. It has the flexibility to deal with the traditional long distance network and also with new and emerging networks where the sources of delay may not be distance related and where the network design approaches do not match older standards and configurations. The 32 in the name reflects its capacity for up to thirty-two 64 kbits/s DS0 channels. This allows the TECO3264 to be designed into both 2.048 Mbits/s E1 systems and 1.544 Mbits/s DS1 (T1) systems. The eight channels unused by DS1 systems still allow economical system implementation. The 64 in the name corresponds to 64 ms of round-trip tail-end delay between the echo canceller and the network reflection source. This is sufficient to permit deployment in the vast majority of networks.

TECO3264 meets the ITU-T G.168 objective tests with considerable margin. More importantly, it meets the subjective expectations of the human listeners in real networks under real call conditions. For non-speech communications, such as voiceband data modem traffic, the modems can turn off the echo canceller functions on a per-call per-channel basis, if desired, using ITU-T recommended 2100 Hz signaling.

The TECO3264 has a large number of parameters and settings, 186 parameters in the read-write registers and 28 parameters in the read-only registers. However, only a handful will be of interest if it becomes necessary to fine tune echo canceller variables for unusual or unique network characteristics in obtaining maximum performance. Settings, such as those for pulse code modulation (PCM) parameters, for μ-law or A-law coding, for μ-law to A-law and A-law to μ-law conversions, and timing alignment are set-and-forget. The recommended set of echo canceller parameter settings are based on years of experience with real networks and accommodate a wide range of normal network conditions. For example, there may be cases where hybrid echo return loss (ERL) is consistently lower than the expected minimum of 6 dB. The minimum threshold return loss may be adjusted to the appropriate level by setting two parameters (ADATA and AVOICE) to optimize performance for the lower ERL range.

The vast majority of parameter settings will never be touched, but will be available, if ever needed, to adjust for different conditions in new wireless land-based or satellite, packet or Internet applications. As refined recommendations for parameter set information are developed for new environments, they will be made available.

The built-in μ-law to A-law and A-law to μ-law conversions of the TECO3264 can be used at the gateways between μ-law and A-law coded networks as a bonus to the basic echo canceller functions.

A very minimal knowledge of the network in which the TECO3264 will be deployed is helpful to obtain maximum benefit of the available performance, but it is not critical for one-time provisioning:

1. Minimum echo return loss (ERL), typically 3 dB to 6 dB, with a 6 dB recommended default if the minimum ERL value is not known.
2. If the network background noise characteristics are known, the noise matching options may be adjusted for an optimum comfort noise match. Otherwise, the recommended settings should prove satisfactory for normal network conditions.

Table of Contents

Contents **Page**

Introduction (continued)

3. If it is known that the signal levels in the given network are consistently higher than normal, there are 6 dB and 12 dB loss pads that may be switched into the receive or send signal paths as desired.
4. Tone disabling options for voiceband data modems and facsimile can be changed if needed.

The TECO3264 microprocessor interface, can be used with either *Motorola** or *Intel*† microprocessors. The microprocessor performs the following external functions for the echo canceller:

1. Static provisioning of the parameters in the read/write Page 0 registers.
2. Normally, the microprocessor does not perform real-time echo canceller calculations except for one possible option. The TECO3264 takes care of all real time functions internally with an option for external call boundary indication to restart echo path cancellation convergence. Normally, the echo canceller detects the start of a new call or a change in the echo path from the signal itself. However, for some networks, such as wireless cellular networks, the external systems have the start (and stop) time information for each call. If desired, this external call boundary information can be fed to the TECO3264 via CALLB (pin 34).
3. Interfacing and monitoring maintenance and error indicators, mostly in the read only Page 1 registers.

Details of the microprocessor/TECO3264 interface are presented in the August 1999 *TECO3264 32-Channel Echo Canceller* Data Sheet (DS99-241PDH).

A fully functional stand-alone evaluation board system, TECO3264EB, is available to allow laboratory and field tests of the TECO3264 for DS1 (T1) and E1 applications. A brief description is found in section Echo Canceller Terms on page 27. See *TECO3264EB TECO3264 Evaluation Board System* User Manual (MN98-072TIC) for detailed information.

* *Motorola* is a registered trademark of Motorola, Inc.
† *Intel* is a registered trademark of Intel Corporation.

Architecture and Functional Description of TECO3264

Description

The TECO3264 32-channel echo canceller device is a 3.3 V, CMOS, very large-scale integration (VLSI) component offering 32 independent channels of echo cancellation. The TECO3264 provides echo cancellation for both DS1 and E1 systems. The device operates from a single 3.3 V supply and requires only an external 8.192 MHz clock and 8 kHz frame sync. Figure 1 shows the basic functions of the TECO3264.

Packaged in a 160-pin, plastic, metric quad flat pack with heat speader (MQFPH) and handling 64 ms of tail-end delay, this device combines high performance, high channel density, low power consumption, high flexibility, excellent maintenance capability, and low cost into a single package.

Built-in, self-test logic affords easy device verification while integrated boundary-scan capability reduces board verification time and cost.

A high-speed microprocessor interface and full user provisionability on device pins provide maximum flexibility.

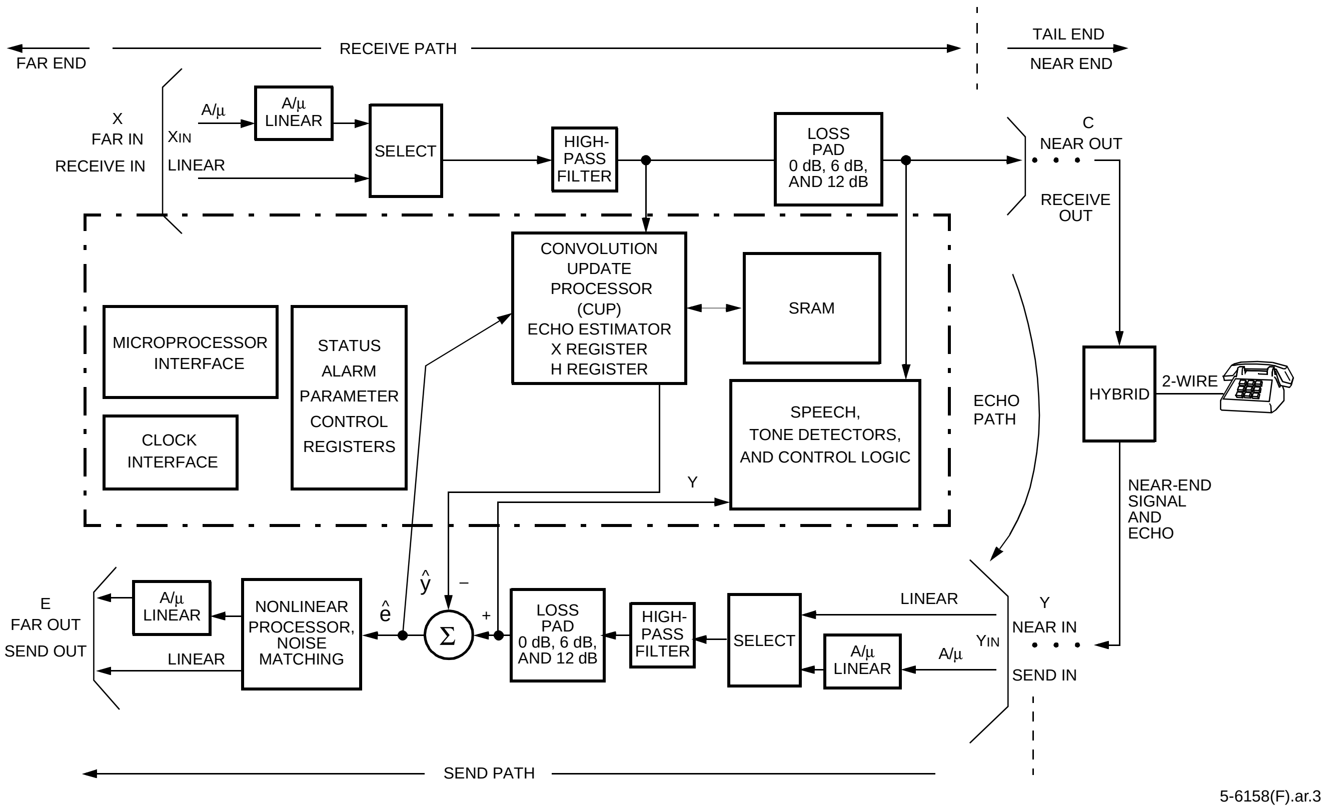


Figure 1. Echo Canceller Functional Diagram

TECO3264 System Interfaces and Formats

PCM Interface

- The TECO3264 provides two independent, serial pulse code modulation (PCM) input ports, and two independent serial PCM output ports. The input ports are labeled X (equivalently far in or receive in) and Y (equivalently near in or send in). The output ports are labeled C (equivalently near out or receive out) and E (equivalently far out or send out), refer to EC diagram in Figure 1, Echo Canceller Functional Diagram, on page 4. Input port X and output port C are the receive path of the echo canceller. Unless there are optional encoding conversions between µ-law, A-law, and linear, the signal passes from X to C unchanged. Input port Y and output port E are the send path and contain the echo canceller subtractor and nonlinear processor functions. Ports X and E face the far end of the circuit. Ports Y and C face the near end of the circuit which may have a 4-wire to 2-wire conversion hybrid circuit. Hybrids are a main source of network electrical echo reflections.
- Each of these serial ports X, C, Y, and E conform to the concentration highway interface (CHI) time-division multiplex (TDM) standard [1].
- The PCM encoding format is inverted or noninverted µ-law, A-law, or 16-bit linear for the X, C, Y, and E ports. The choice between linear and companded encoding is provisionable per port independently. The choice of companded encoding law is provisionable for all-channel[1] for each port independently. Most applications use inverted µ-law or A-law coding.
- The serial-bit rate on all ports is either 4.096 Mbits/s with thirty two 16-bit time-slots, or 2.048 Mbits/s with thirty two 8-bit time slots. The selection of the I/O bit rate applies to all four ports. Mixing of the 2.048 Mbits/s and 4.096 Mbits/s is not supported. The linear PCM format is not supported for the 2.048 Mbits/s mode.
- The TECO3264 provides parity checking on the X and Y PCM input streams. The parity is programmable for odd, even, or none at all. Parity can be calculated over the 8-bit PCM word and placed in the ninth bit or, alternately, over the 15-bit word and placed in the sixteenth bit. Parity is not supported when the 16-bit linear PCM encoding format is used. The parity options are provisioned independently for each port.
- The TECO3264 provides parity generation on the E and C PCM output streams. The generated parity is programmable for odd, even, or none at all. Parity can be calculated over the 8-bit PCM word and placed in the ninth bit or, alternately, over the 15-bit word and placed in the sixteenth bit. Parity is not supported when the linear PCM encoding format is used. The parity options are provisioned independently for each port.
- In the 4.096 Mbits/s companded mode, bit 8—bit 15 will pass through unaffected unless the parity generation option is enabled on the output port. In that case, parity for bit 0—bit 7 will overwrite bit 8 or bit 15 depending on the parity option chosen.
- Conversion from either A-law to µ-law or µ-law to A-law meets ITU-T conversion requirements [2].
- An alternate (adjusted) µ-law to A-law or A-law to µ-law conversion is provided [2]. Adjusted conversions preserve 7 of 8 encoded bits for tandom µ-law to A-law to µ-law or A-law to µ-law to A-law conversions. This feature is provisionable on an all-channel basis.

1. Digroup often is used to refer to all 24 channels for a 1.544 Mbits/s DS1 application. In DS1 applications, the term group means 12 channels or DS0 time slots. This document uses the term all-channel to refer to all 24 channels for DS1 or all 32 channels for E1.

TECO3264 Functions

Linear Processing

Convolution

The convolution processor (CUP) provides for tail-end delay of up to 64 ms (512 taps). Tail-end delay is the round trip delay experienced by a signal exiting the C port, being reflected, and then entering the Y port.

Gain Normalization

- Gain normalization is based on a measure of Xin power. It is provided to make convergence time independent of X input signal level. The prenormalized gain is provisionable on an all-channel basis and provided by the convolution gain control circuit.
- The set of peak factors available for use in the gain normalization process are provisionable on an all-channel basis. Various indications are used to select the proper peak factor under a given set of conditions (refer to the Control section on page 8).

High-Pass Filter

- All-channel selectable high-pass filter (≥10 Hz) is provided for X input signal.
- All-channel selectable high-pass filter (≥10 Hz) is provided for Y input signal.
- The high-pass filters are disabled on a per-channel basis whenever a channel is either in 64-clear or EC_DISABLE mode.

These high-pass filters block any direct current present on the echo canceller ports from reaching the echo canceller circuitry.

Nonlinear Processing

Even with the 34 dB echo return loss enhancement, residual echo $\hat{e}$, called ehat, can be occasionally heard on low-noise trunks. In order to reduce this effect, a nonlinear processor (NLP) inserts either loss (finite or infinite), or a controlled level of noise in the echo return path when there is no double talking, i.e., only far-end speech is present. The 34 dB echo return loss enhancement approaches the maximum possible with µ-law and A-law companded signals. During periods when near-end speech is detected, the NLP and comfort noise are removed immediately.

Thresholds for determining when only far-end speech is present are set to eliminate clipping of near-end speech. The thresholds depend on the relative level difference between near-end and far-end signals. These thresholds are defined as SBETA and FBETA. The NLP threshold in dB is given by:

NLP Threshold (in dB) = $20 \log_{10} (B/8192)$, where B is equal to either SBETA or FBETA.

There is also a transient echo control option designed to remove echo during the initial start-up of echo canceller convergence. This is engaged whenever the power measure of Eout greatly exceeds power measure of Yin, a situation that can be only a transient one.

TECO3264 Functions (continued)

Linear Processing (continued)

Noise Matching

Noise matching is provisioned all-channel and used to alleviate the effects of the NLP switching in and out as heard by far-end talkers when they start and stop talking. When selected, it automatically measures the near-in noise level and instead of inserting infinite loss in the echo return path, inserts an equivalent amount of noise. This inserted noise is sometimes called comfort noise. From experience, customers expect to hear some low-level noise during the gaps in speech and may think a totally quiet line has been disconnected.

Soft Nonlinear Processor

- The soft nonlinear processor (SNLP) mode allows a smoother and slower transition from NLP disabled (switch closed) to NLP enabled with noise matching. An abrupt change in background noise is noticeable.
- The SNLP mode is provisionable on an all-channel basis.

Noise Transparency

- The transparency mode allows the noise-matching data insertion to be colored by low-level, near-end input samples intended to improve spectral noise-matching performance. The inserted comfort noise sounds more like the real background noise that it is replacing and the change is less noticeable to the customer.
- The transparency mode is provisionable on an all-channel basis.

TECO3264 Functions (continued)

Control

Far-End Speech Detector (FES)

Two measures of far-end energy (i.e., signals coming in the X port) are calculated. The first is signal power computed from all the samples stored in the X register. The second is based on the peak signal magnitude stored in the X register. This measure is adjusted by a peak factor that depends on the spectral (narrowband vs. wideband) content of the signal and state of the canceller.

- Far-end speech (FES) is declared (FES = TRUE) if either of these measures exceeds a threshold value. Note that there is no hangover count involved.
- The peak factors are provisionable on an all-channel basis.
- The threshold value is provisionable on an all-channel basis.

Near-End Speech Detector (NES)

The TECO3264 system provides two independent indications that near-end speech (signals coming in from the 2-wire circuit via the hybrid and into the Y port) is present. These indications are combined with other indications to control convergence.

- The first indication of near-end speech is derived by comparing the X-input power measure to Y-input power measure. When the Y-input power measure exceeds the X-input power measure by a provisionable threshold value, near-end speech is declared. The H-register values are frozen during NES = TRUE. The H-register values may adapt during FES = TRUE and NES = FALSE.
- The threshold used in the first indication should be selectable from one of two all-channel provisionable values. The means to select between the two values will be the narrowband energy (NBE) detector state or, alternately, an external voiceband signal classifier (VSC) input indication. The voiceband signal classifier with the narrowband energy detection decides whether a wideband signal, such as speech, or a narrowband signal, such as a tone, is present. An external pin is provided for the VSC serial control input.
- The second near-end speech indication is provisionable on an all-channel basis.

Narrowband Energy Detector (NBE)

- The narrowband energy detector is used to classify the far-end X input signal as wide- or narrowband.
- A means is provided to substitute the internal NBE indication with an external VSC indication. This feature allows the use of a more accurate voice/data classification of the signal by an external processor.

TECO3264 Functions (continued)

Control (continued)

Tone Disabling

The echo canceller is equipped with two independent, tunable tone disabler functions (TD1 and TD2) to disable the echo canceller when voiceband modem, facsimile data, or other specified tone signals are transmitted through the canceller.

- The tone disabler characteristics conform with ITU-T Recommendations G.164 [3], G.165 [4], and G.168 [5]. Tone detector TD1 is usually used.
- The tone disabler is used to disable the echo canceller on the detection of specified tones, but does not disable on speech.
- The tone disabler provides tone disabling function for 32 channels in send (Y-E), and 32 channels in receive (X-C) directions.
- The tone disabler responds to a disabling signal that may be present in the send or receive direction.
- The tone disabler detection circuit is capable of detecting a 2100 Hz ± 21 Hz tone, and disabling the echo canceller as specified in ITU-T Recommendation G.164 [3].
- For correct operation of V-series modems, the tone disabler detection circuit is capable of detecting a 2100 Hz ± 21 Hz tone with periodic phase reversals inserted in that tone.
- The TD2 tone disabler is capable of detecting tones in the frequency range of 2000 Hz to 2110 Hz tone for voice path assurance (VPA). VPA is an end-to-end path continuity test that may be performed by the network before the circuit is given to the customer for use.
- VPA detection is provided on the receive path.

H-Register Freeze

The H-register contains the echo canceller's model of the tail-end echo path. Each of the 32 E1 or 24 DS1 channels has its own echo model calculated for each customer call. The model adapts for changes during the call.

Control of the H-register freeze function can come from any of the following sources:

1. External global control—an external pin is provided to globally freeze the H-register for all channels. A mode is provided to allow per-channel serial control of the freeze function.
2. NES/FES control—the NES and FES speech detectors are combined to determine when to freeze. The freeze function is enabled whenever NES is present or FES is not present.
3. NBE control—NBE detector is used to freeze the H-register. The freeze function is enabled whenever NBE is present.
4. There is a per-channel freeze control via microprocessor interface.

H-Register Reset

The H-register reset (setting of all H-register taps to zero) function is performed on a per-channel basis. The sources that can affect a reset are as follows:

1. External control—a means is provided through the microprocessor interface to reset the H-register on a per-channel basis.
2. In addition to the microprocessor interface control, an input pin is provided to globally reset the H-registers for all channels. A mode is provided to also allow per-channel serial control of the H-register reset function.

TECO3264 Functions (continued)

Control (continued)

Nonlinear Processor Control

The NLP feature in the TECO3264 system is controlled by several sources. The following is a list of these sources:

1. Per-channel NLP disable—an external pin is provided to allow per-channel serial control of the NLP.
2. Residual echo $\hat{e}$ vs. X calculation—this is the dominant means of NLP control. If the power measure on X (after coprocessor) exceeds the power measure on $\hat{e}$ by a certain provisionable threshold value, the NLP is operated.
3. The differential sensitivity threshold used in the $\hat{e}$ vs. X calculation has the capability of taking on two independently provisionable values depending on the convergence mode (refer to the Convolution Gain Control: Fast and Slow Convergence Modes section on page 10).
4. Per-channel microprocessor control—the NLP can be controlled via the microprocessor interface on a per-channel basis. This control either allows real time control of the NLP by the echo canceller or turns the NLP on full time.
5. A mode is provided to freeze the state of the NLP whenever the H-register is frozen. This option is provisioned on an all-channel basis via the microprocessor interface.

Convolution Gain Control: Fast and Slow Convergence Modes

The TECO3264 system is equipped with a convolution gain control circuit whose purpose is to accelerate the convergence rate under certain conditions on any given channel. This is called the fast convergence mode. In conjunction with this feature, the system provides several means for detecting when the echo canceller (EC) is not converged. The specific measures for EC convergence are done on a per-channel basis and are listed below:

1. H-register reset—when the EC transitions out of the H-register reset state (for any reason), the EC may be considered not converged.
2. External control—the EC provides a means for an external (to the TECO3264 system) source to indicate that the EC is not converged. This indication is communicated via the microprocessor interface. In addition, an external pin is provided for this function.
3. $\hat{e}$ vs. Y calculation—a parameterized calculation determines if the EC is converged by examining the power measure at $\hat{e}$ vs. the power measure at Y. If the power measure at $\hat{e}$ is sufficiently larger when compared to the power measure at Y, the EC is considered not to be converged.
4. VPA detected to not detected transition—since VPA provides an indication of call setup, a VPA detected to not detected transition on the VPA detector indicates that the EC is not converged.
5. Each of the aforementioned indications are optionally enabled or disabled on an all-channel basis. (Action takes place on a per-channel per-call basis.)
6. When the circuit is recognized as not converged by one of the EC convergence indications, the convolution gain control circuit may apply an alternate, provisionable convolution gain for a provisionable amount of adaptation time.
7. When the circuit is recognized as not converged by one of the EC convergence indications, the convolution gain control circuit may apply an alternate, provisionable NLP differential sensitivity threshold for a provisionable amount of adaptation time.
8. When the circuit is recognized as not converged by one of the EC convergence indications, the convolution gain control circuit may apply an alternate, provisionable peak factor value to the gain normalization circuit for a provisionable amount of adaptation time.

TECO3264 Functions (continued)

Control (continued)

Control of Channel Processing Features

- The TECO3264 system provides a per-channel 64-clear channel or bypass option. When enabled on a given channel, the 8-bit data words for that channel will pass through unaffected. This occurs in both directions of transmission and overrides all other channel control options. The eight unused channels in DS1 applications may be set to 64-clear mode.
- The basic EC (convolution processing, NLP, and high-pass filter (HPF)) function can be controlled on a per-channel basis. When disabled on a given channel, and the channel has not been placed in the 64-clear mode, A-law to μ-law or μ-law to A-law conversion can still take place if provisioned.
- The basic EC function can be enabled or disabled on a per-channel basis via the microprocessor interface.
- The basic EC function can be enabled or disabled per-channel via an external pin (ECDIS, pin 26).
- The basic EC function can be controlled via in-band tone disabling (refer to Tone Disabling section on page 9). The echo canceller function may be optionally disabled with 2100 Hz tone detection with phase reversal, 2100 Hz tone detection with or without phase reversal, or not at all.
- The echo cancelling function can be controlled via in-band VPA tone disabling (refer to Tone Disabling section on page 9).
- The 64-clear function can be enabled or disabled on a per-channel basis via the microprocessor interface.
- The 64-clear function can be controlled via in-band tone disabling (refer to Tone Disabling section on page 9). The 64-clear function may be optionally enabled with 2100 Hz tone detection with phase reversal, 2100 Hz tone detection with or without phase reversal, or not at all.

  Note: A provisionable minimum signal power must remain on the channel to maintain the 64-clear function.
- The frame delay through the device in each direction is the same for all modes of operation.

Echo Canceller Evaluation Board System (TECO3264EB)

To allow laboratory and field tests of the capabilities and performance of the TECO3264 integrated circuit, Lucent has developed a fully functional, stand-alone echo canceller evaluation system for T1 (DS1) and E1 applications. See Figure 2, TECO3264EB Hardware Block Diagram on page 13.

A Lucent Technologies Microelectronics Group T7630 Dual T1/E1 Terminator performs the E1 or T1 line interface and framing functions for the TECO3264. PCM data and clock information exchange between the TECO3264 and T7630 is by a programmable 2.048 MHz or 4.096 MHz serial time-division multiplex (TDM) bus, called the concentration highway interface (CHI)[4]. The CHI input/output is connected to four Lucent T7270 time-slot interchange (TSI) circuits configured as a four-channel time-slot channel network switch. The external control interface to the T7630 is by an RS232 link to an onboard *Intel* 87C51FC microprocessor. The external control interface to the TECO3264 is by an RS232 link to an onboard *Motorola* MC68360 microprocessor.

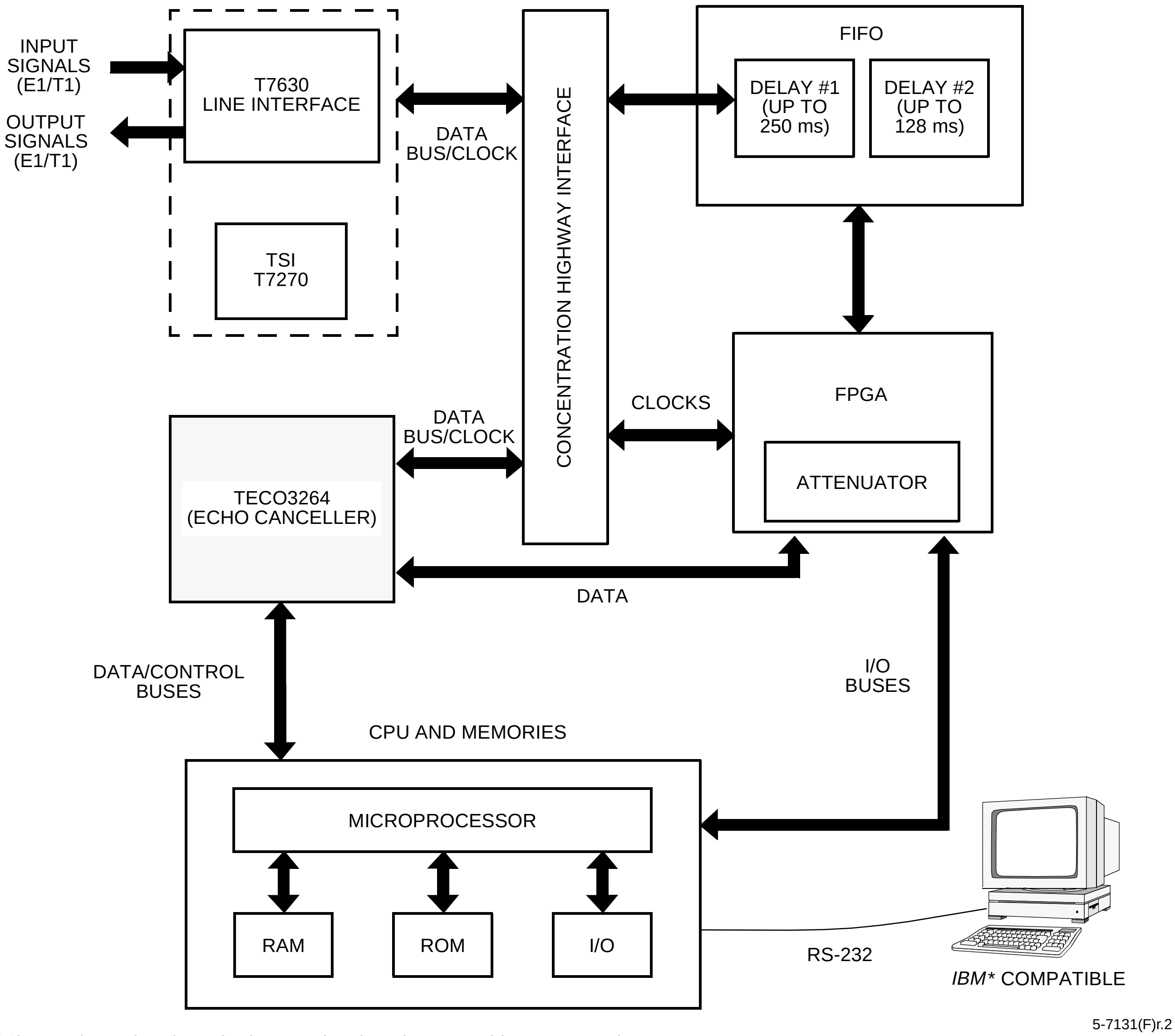
The external equipment required to operate the TECO3264EB system consists of sources of 3.3 Vdc at 1.0 A and 5.0 Vdc at 3.0 A and an *IBM** compatible computer to run the ANSI terminal interface user-interface software. The user-interface software for the TECO3264 and T7630 is supplied. External T1 or E1 communications test analyzer systems or live network circuits may be connected to the line input and output ports.

The system can be connected to external telephone line circuits to test with live hybrid echo, attenuation and tail-end delay.

Full details on the TECO3264EB system are found in the *TECO3264EB TECO3264 Evaluation Board System User Manual* (MN98-072TIC).

* *IBM* is a registered trademark of International Business Machines Corporation.

Echo Canceller Evaluation Board System (TECO3264EB) (continued)



* *IBM* is a registered trademark of International Business Machines Corporation.

Figure 2. TECO3264EB Hardware Block Diagram

Echo Canceller Applications

Wired Telephone Network Applications: Local to Long Haul

To most telephone customers the worldwide telephone system is a big box whose only visible parts are the phones on their desks and sometimes some cables hung on poles throughout the neighborhood. One picks up the phone handset, dials a number, magic happens, and one can speak with someone across town or halfway around the world. Your voice goes in here and comes out over there. Even some technically minded customers have little understanding of the many parts of the network that are involved. This section will present some basic sketches and explanations as background to where voice frequency echo cancellers, such as the TECO3264, fit in the overall picture. Other subsections and paragraphs deal with individual topics in a little more depth and detail.

A Local POTS Customer Connection

Figure 3 below depicts the local connection for a typical plain old telephone service (POTS) customer. On the customer premises, the most visible piece is the telephone station set. In addition to station sets, there may be data modems, facsimile (FAX) machines, answering machines, and caller identification boxes connected in parallel. The inside wiring with two-wires starts at the interface to the network two-wire loop at the entrance to the building, in a basement, garage, or other convenient point. The inside wiring runs about the premises and appears for a connection wherever the customer desires. The network cable, called the loop, connects from the customer premises to the local telephone building central office (CO). The loop connects via distribution frames and inside CO wiring to a POTS line unit in the switching machine. The switch here is a digital switch. The local digital switch is a very specialized, very reliable computer controlling four-wire connections between other local customers or long-haul connections. Since internally the switch is a four-wire system, each POTS customer port has a four-wire to two-wire conversion circuit called a hybrid. In North America, digital loop carrier (DLC) systems are bringing the line unit and hybrid functions out to within 2 miles or less of the customers who are at long distances from the CO.

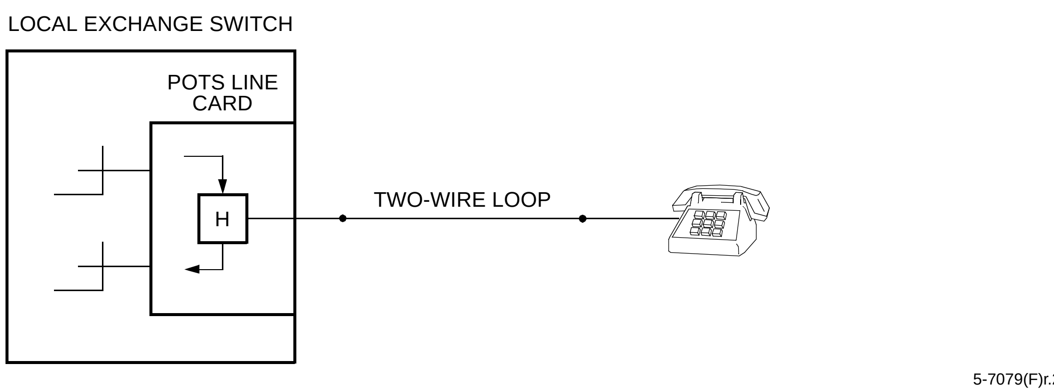


Figure 3. POTS Line Card and Telephone Switch

Echo Canceller Applications (continued)

Wired Telephone Network Applications: Local to Long Haul (continued)

Local Loop Detail

Figure 4 below shows a little more inside the telephone and the switch line unit. The telephone has its own hybrid to convert from the two-wire handset speaker and two-wire microphone to the two-wire network loop. This hybrid and its balance network is designed for a deliberate mismatch of about 10 dB to provide some feedback, called sidetone, from the microphone to the speaker. This sidetone helps the customer to unknowingly speak at a level that is about right for the network. The two-wire metallic loop connects to the local switch (or to a digital carrier system). The switch POTS line unit performs the hybrid two-wire to four-wire conversion (discussed in more detail later), analog-to-digital and digital-to-analog transmission conversion, and also supplies the interface for the CO battery to power the telephone sets down the loop and apply a ringing generator to ring the customer phone for an incoming call. The balance networks, called compromise balance networks (CBNs), for the line unit hybrid are designed to match loop input impedance well enough for low-loss switch operation.

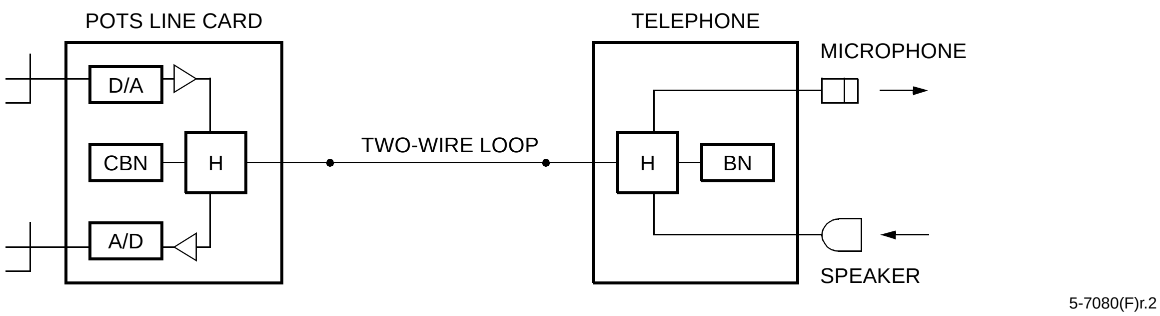


Figure 4. POTS Line Card and Telephone Set

Echo Canceller Applications (continued)

Wired Telephone Network Applications: Local to Long Haul (continued)

Local POTS Call

Figure 5 below depicts a local POTS call with two customers connected by the switch. North American loop design rules limit the maximum loss at 1 kHz to about 9 dB with the average loop having about 4 dB loss at 1 kHz. Loss across the switch is normally close to 0 dB. So the customer to customer loss at 1 kHz can be as much as 18 dB and averages about 8 dB. Too little loss results in too high levels of speech so a modern switch can insert loss to maintain a minimum loss. If one counts the hybrids in the station sets and line units there are at least four hybrids in the circuit, none of which are perfectly balanced. But with the short distances and low delays, echo is not perceived and is not a concern.

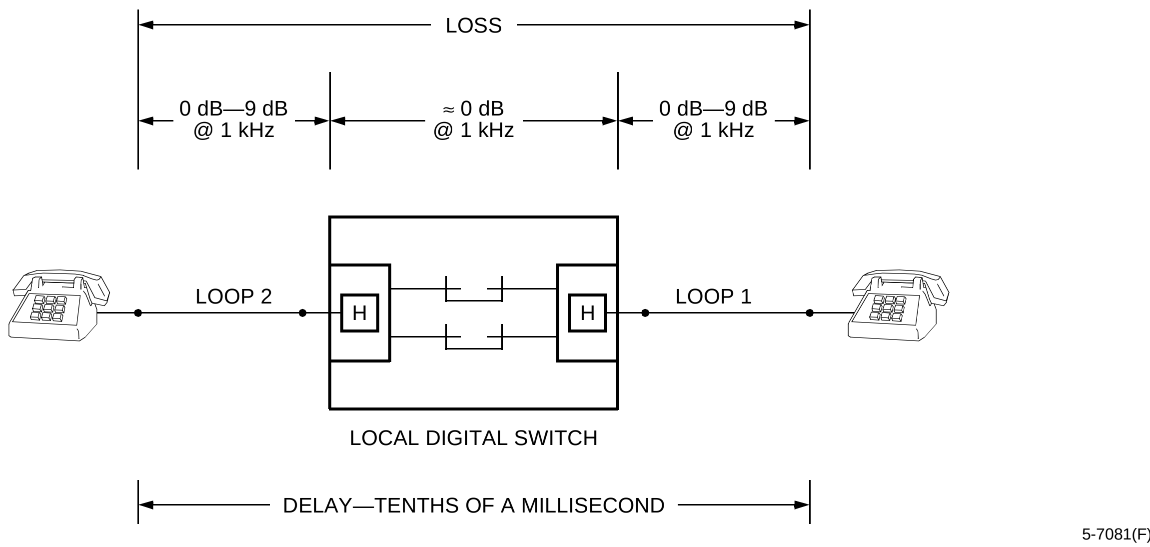


Figure 5. Local POTS Call Connection

Echo Canceller Applications (continued)

Wired Telephone Network Applications: Local to Long Haul (continued)

A Long Distance Call Connection

Figure 6 below shows an end-to-end long distance call connection starting with a POTS customer and local switch at each end. The connection is symmetric about the middle. The local switches connect to trunk facilities that connect to the toll switches. On the other side of the toll switches, there is an echo canceller oriented for cancelling echoes that come from the local loop hybrid. This position allows the echo canceller to be shared among many trunks from the local network on a call-by-call basis. The tail-end delay is mainly the signal propagation delay from the echo canceller receive-out/C port through the toll switch, the toll-to-local trunk facilities, and through the hybrid and back to the echo canceller send-in/Y port. With no speech compression, packet switching, or forward error correction, the switches and other transmission terminal gear contribute little to the delay. The 64 ms tail-end delay capacity of the TECO3264 will account for the greater majority of such local to toll connections.

Note that the long-haul delay inside the toll network is not addressed by the echo cancellers, even though it can be hundreds of milliseconds if geosynchronous satellite links are included. The echo cancellers are intended to prevent the echoes from the hybrids from getting back to the long delay links.

More and more intercontinental voice (and data) traffic is being moved to an ever growing network of undersea optical fiber cables where the delays are much shorter (but not negligible) and the bandwidths are much greater than the radio links through satellites. Wavelength division multiplex techniques are allowing many light signals in one fiber with terabits of digital data transport capacity. Such optical undersea cables are nearing deployment with all but Antarctica on the schedule. While the occurence of long 100 ms delays may decrease, the long-haul delays just due to terrestial distances will still be enough to require echo control.

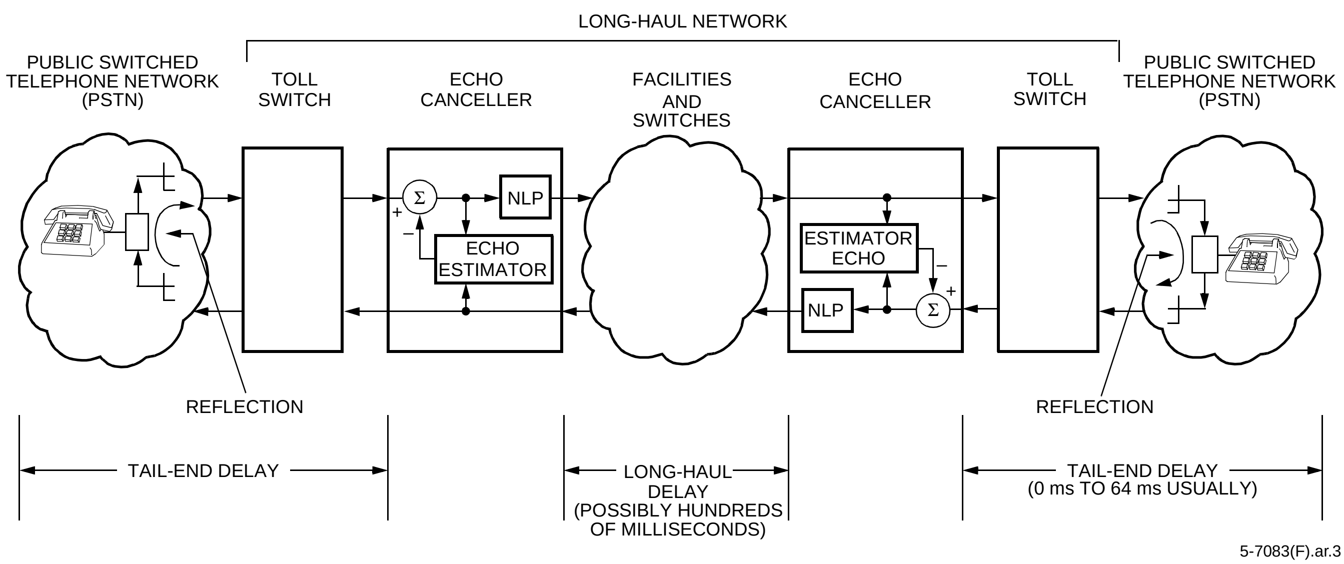


Figure 6. Long Distance Telephone Connection

Echo Canceller Applications (continued)

Wired Telephone Network Applications: Local to Long Haul (continued)

Cancelling on the Edge, or Maybe Not

When a major long-haul toll network provider connects to the PSTN with a major local exchange provider, there is usually enough of a cross section of trunks that DS1 or E1 based network echo cancellers with 24 or 32 channels per card can be placed at the edge of the toll network behind the toll switch. This avoids picking up the internal long-haul delay from the toll network and lets echo canceller systems with a 64 ms tail-end delay capacity, like systems based on the TECO3264 ASIC, handle virtually all of the echo control duties. The local exchange carrier concentrates their long-haul connecting traffic so that the number of connections to the long-haul carriers can be handled in as few locations as practical.

In several countries, like the U.S.A., the governments are allowing competition to the current local exchange providers by new telephony service providers. Most of these new service providers are small and may have only a few circuits in any one city. When these circuits involve long distance connections, the interconnections have to be carried farther back inside their long-haul network to pick up echo control. A result is that some of the long-haul delay may be included in the tail-end delay seen by the echo cancellers. Instances of tail-end delays greater than 64 ms and as great as 150 ms, in a very small number of cases, have been encountered. Such cases are expected to continue to be in the minority, and there are means to deal with them. If it is known that the delay for a given trunk group is always greater than 64 ms, a fixed delay can be added externally to the TECO3264 chip. The potential market is being studied for a new version of TECO that can handle longer-end delays. When cross sections of trunks having less than approximately six circuits need more echo control, DSP-based solutions using algorithms similar to those built into the TECO ASIC may be considered. When there are ten or more trunks, using the TECO3264 with external delay padding may be economical, even when most of the channel capacity is not used. The long-end delay and small cross-section application situation is a matter of study to develop the best technical and economic approaches.

Echo Canceller Applications (continued)

Wired Telephone Network Applications: Local to Long Haul (continued)

Echo Cancellers in Nontraditional Networks, Connections to PSTN

Many of the new voice transport schemes and networks such as digital cellular phones, PCS, wireless local loops, low-earth orbit satellites, speech compression, forward-error correction, and packet switching create transmission delay that is not distance related. Figure 7 shows only the end points, delay, and echo control of these networks. The delays in a hybrid fiber coax (HFC) telephony system will still be mainly due to distance (unless very low bit-rate voice or packet transmission is used). When networks interconnect to the traditional PSTN, they can experience echo that results from the delay inside their networks and from the hybrids in the local POTS networks. A solution is to install network electrical echo cancellers in these networks at the gateways to the PSTN. The echo cancellers are pointed at the PSTN local hybrids.

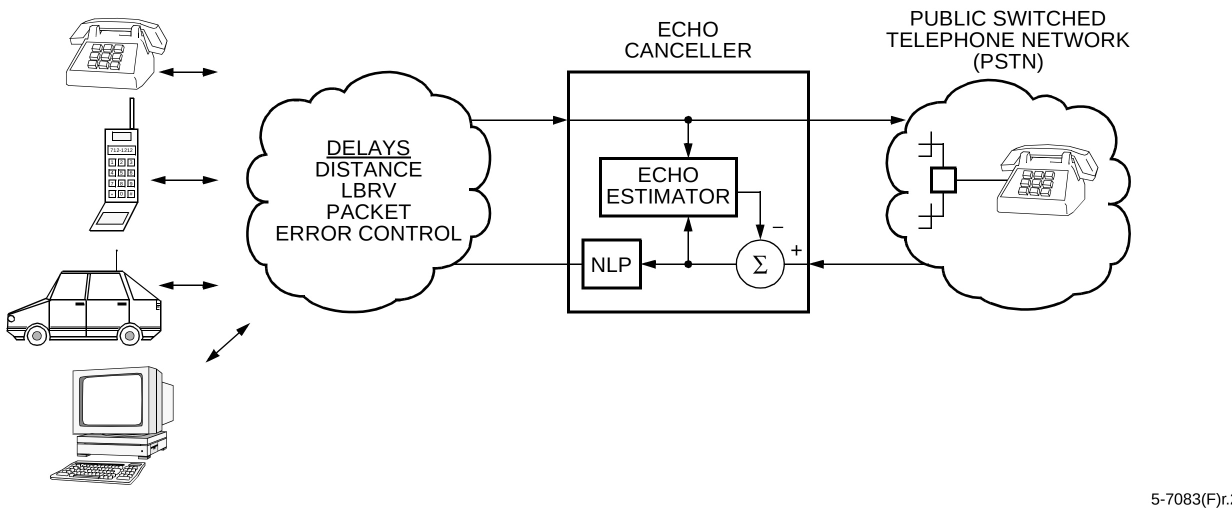


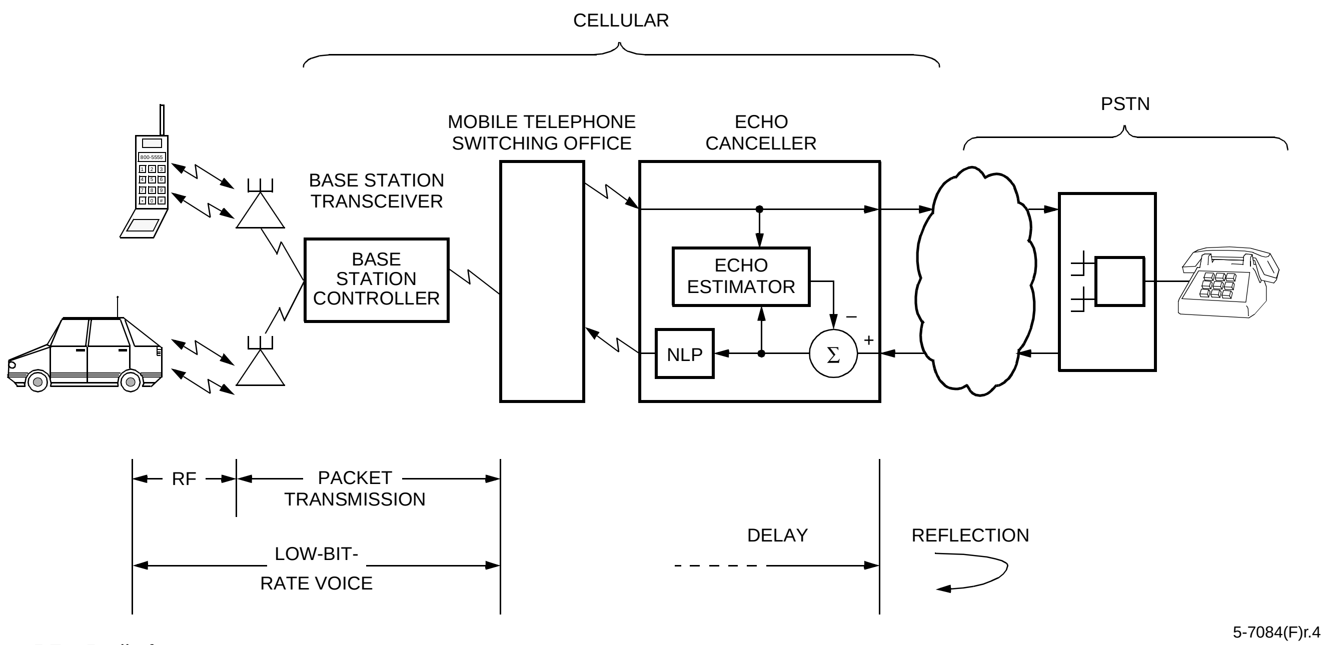
Figure 7. New Telephony Transport Network

Echo Canceller Applications (continued)

Digital Wireless, Cellular Low-Bit-Rate Voice

Figure 8 below depicts a simplified view of a digital cellular wireless system. Digital wireless telephony systems developers found themselves in a situation like the early digital landline systems, where every bit of data, every bit per second, every Hz of analog bandwidth cost circuit complexity and money. The very robust 64 kbits/s µ-law and A-law schemes were too much when, hopefully, only one or two analog-digital-analog conversions would be needed [2]. First, a standard 32 kbits/s low-bit rate voice (LBRV) scheme of nearly the same quality as 64 kbits/s was developed, just to double the voice channel capacity. Then, 16 kbits/s schemes were developed, then 8 kbits/s schemes. Schemes with only 2 kbits/s, or even less for voice communications, are under study. The GSM wireless system, which is becoming dominant outside the U.S.A., often uses 13 kbits/s voice coding. Reportedly, customers can be happy with the results. The catch is that these lower and lower bit-rate voice coding systems require a lot of digital signal processing calculations to implement the many complex algorithms needed to yield acceptable speech for human ear. These calculations take time, and time equals delay in the voice path. The delay increases rapidly as the bit rate goes down. Milliseconds of delay appear in calls where the old 64 kbits/s approach the conversions are essentially instantaneous. Echo cancellers can be used. The vocoder functions may be in either the base station controllers or in the mobile telephone switching office. (For general information on speech coding, see reference 6.)

Another aspect of digital cellular systems is that packet transmission is often used to transport voice packet, call control signaling, and systems control between the wireless base stations in the field and the mobile switch. Long or variable frame packet schemes like frame relay can introduce considerable delay (see Packet Transmission Impairments section on page 26).



Note: RF = Radio frequency.

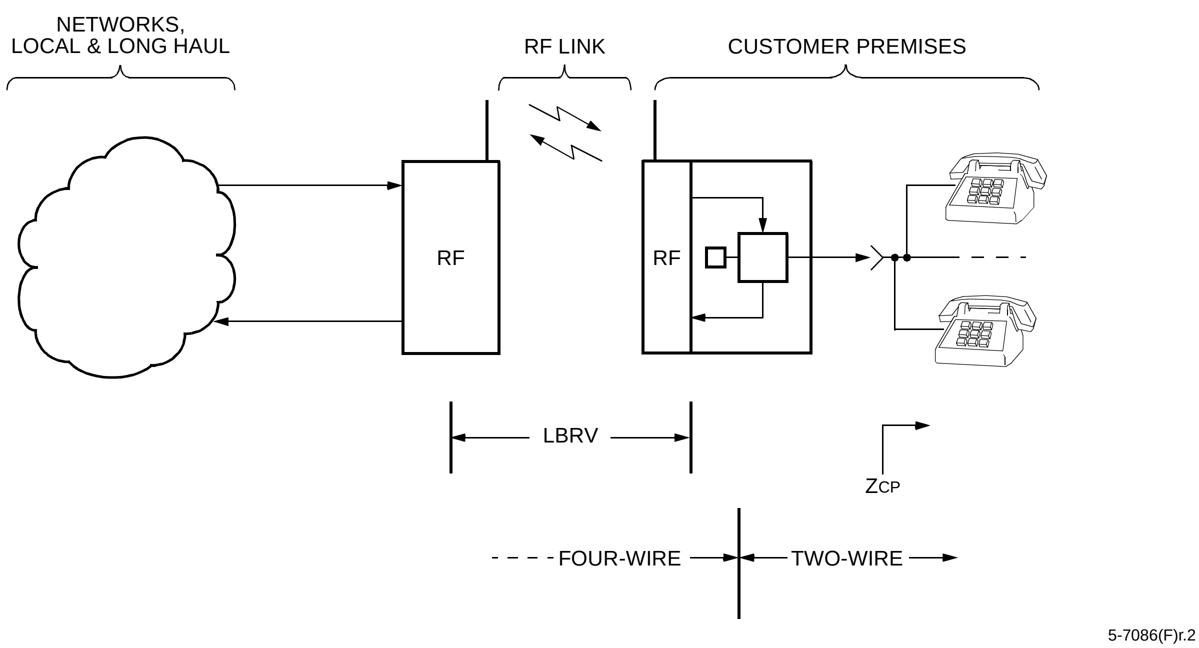
Figure 8. Cellular Telephone

Echo Canceller Applications (continued)

Wireless Local Loops

The International Telecommunications Union (ITU) predicts that by 2002, seventy-five percent of newly installed POTS customer loops will be wireless, not copper twisted pairs (see Figure 9 below). This will likely be in rural areas in developing countries that do not have the huge embedded base of wire that developed countries have. Wireless technologies promise the possibility of reaching customers with voice and data services without installing copper wire to every home. There is speculation that there may not be enough copper available to deploy in China. Links from the remote rural areas may well be wireless via earthbound or satellite facilities as well.

These wireless local loops will bring the four-wire-two-wire conversion all the way out to the hybrids in the customer station sets. As shown in the figure, the terminal equipment box for the wireless loop will contain a hybrid to convert from the four-wire radio link to interface with the two-wire circuits in the customer premises. The choice of the hybrid compromise balance network (see the hybrid discussions below) will much more of a compromise than the balance networks used at the central office or in digital loop carriers. From the central office the hybrids see the loop input impedance which, except for very short loops, is determined mainly by the cable characteristics with the customer premises input impedance buffered by the loop loss. While individual loop input impedances at the central office vary considerably from one another, they are similar enough for the simple compromise balance networks to control stability and singing in the switch. For the wireless terminal hybrids the customer input impedance will be determined mainly by the various customer devices that may by connected in parallel and off-hook at the same time: telephone sets, answering machines, caller ID, fascimile, and modems, etc., with the short cable runs having little effect. The customer input impedance can vary greatly even during the call as devices go off- and on-hook.



Note: RF = radio frequency.

Figure 9. Wireless Local Loop

The hybrid balance network will probably have to be the standard 600 Ω + 2 µF network (in the U.S.A.) used for customer premises switches or a network developed to look somewhat like the impedance of telephone sets. The electrical echo as seen by the electrical echo cancellers back in the network will vary accordingly. As discussed in the next paragraph, this electrical echo may be highly nonlinear as a result of the nonlinear speech coding in the wireless links.

Echo Canceller Applications (continued)

Wireless Local Loops (continued)

Some wireless local loop systems plan to use the standard low-bit-rate voice (LBRV) 32 kbits/s adaptive delta PCM (ADPCM) voice encoding instead of 64 kbits/s µ-law or A-law encoded PCM to save radio bandwidth. The 32 kbits/s ADPCM is very close to the 64 kbits/s companding in terms of near toll quality and linearity. Some of the wireless local loop systems plan to use technology and speech coding developed for digital cellular and personal communications systems (PCS) systems. These may use speech compression techniques such as linear predictive vocoders at 13 kbits/s, 8 kbits/s, or 5.6 kbits/s to transport voice. Vocoders at 4 kbits/s are in the lab. In spite of linear in the name, these devices use highly nonlinear encoding and decoding approaches to achieve ultra low rates. The output of a vocoder is not a compressed version of not-quite-but-very-nearly-linear µ-law or A-law PCM, but is a set of instructions on how to build a signal at the decoder that sounds like the original speech to the human ear and brain using white noise and a library of sound samples that can be scaled in frequency and amplitude. One problem with the ultra low rate vocoders is developing algorithms that will work for most human languages. Variable bit-rate voice transmission is another scheme that may be used where the bit rate varies with time as the speaker is talking or silent.

Network electrical echo cancellers, like the TECO3264, depend on the echo being no more nonlinear relative to the original signal than µ-law or A-law, or about 34 dB. As more nonlinearity is added to the circuit, the maximum available ERLE will degrade accordingly.

With the more variable electrical echo due to the more variable customer premises hybrid reflections that pass through nonlinear vocoders, the challenge to control echo in the network will be greater.

Acoustic echoes with handsfree speakerphone home phones on wireless loops or in cell phones are handled by acoustic echo cancellers at the analog stages of the cell phone before encoding to digital. Residual acoustic echo can be reduced by adjusting TECO3264 parameters.

Echo Canceller Applications (continued)

Hybrid Fiber Coax (HFC) Telephony

Figure 10 depicts a simplified view of a hybrid fiber coax (HFC) telephony connection. Here is another use of the word hybrid to mean a mix of transmission media, in this case metallic coaxial cable of the type used to carry television signals in electrical form and optical fiber for signals in light form.

The cable television industry has been working for years to develop technology to transport telephone communications on their wideband coax and fiber facilities along with the television channels. If they could do that, they could support POTS-like telephone service and tap into the huge telephone system revenues, and bypass the traditional local telephone companies. For various reasons, including technical, political, reliability, funding concerns, HFC telephony has not caught on in any major way. As of June 24, 1998, that may well change. The largest long distance telephone network, announced it is acquiring the second largest CATV system, for $48 billion. With this size of financial commitment, HFC telephony could well turn out to be a major part of voice telephone communications in the years to come.

As in the case of wireless local loops, the four-wire portion of an HFC telephone circuit will extend the four-wire network right up the customer premises interface where the HFC hybrid will do the conversion. As discussed for wireless loops, the customer input impedance is much less controlled than the impedance seen at the central office hybrids and will result in probably lower and more variable ERLs as seen by the network electrical echo cancellers. In contrast to the wireless local loops, the full 64 kbits/s µ-law or A-law companded signals may be transported all the way to the HFC terminal hybrid at the customer premises. If this is so, HFC telephony won't have the nonlinearity concerns of the wireless systems using low-bit-rate voice vocoders.

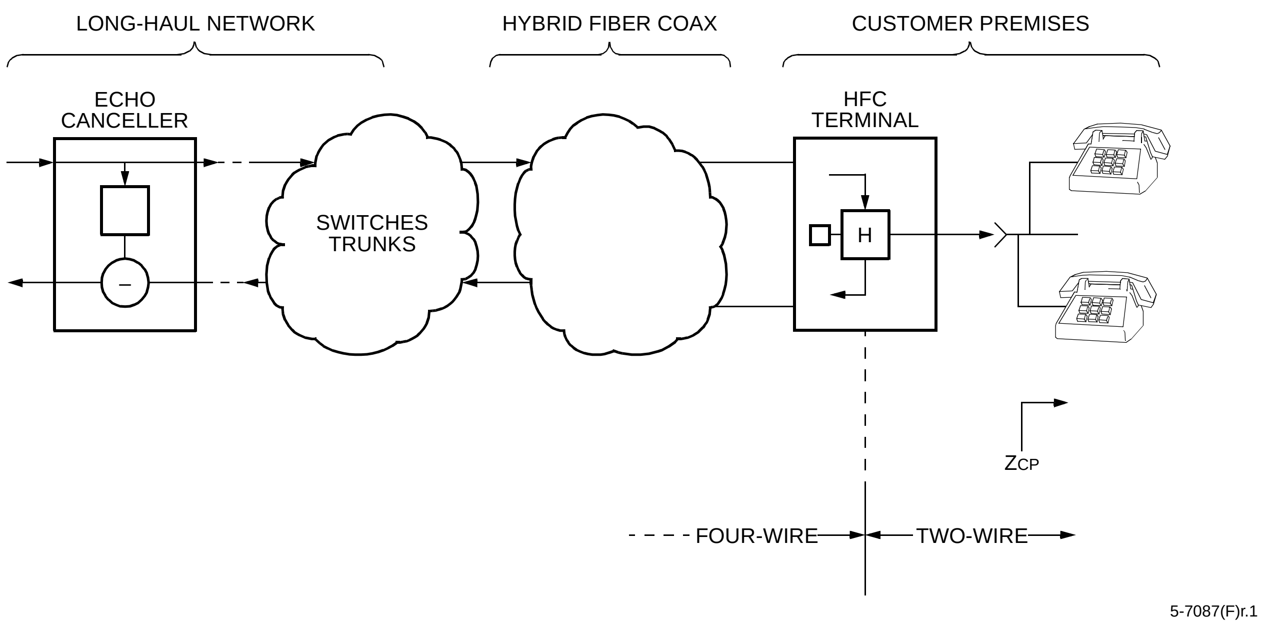


Figure 10. Hybrid Fiber Coax (HFC) Telephony

Echo Canceller Applications (continued)

Integrated Services Digital Networks (ISDN) Basic Access, The New Digital POTS

ISDN basic access (also called basic rate access (BRA)) has for many years been touted as the digital replacement for analog POTS. ISDN BRA provides a net digital payload to the customer premises of 144 kbits/s. This 144 kbits/s is usually divided into two 64 kbits/s bidirectional channels labeled B (bearer) and a 16 kbits/s channel labeled D (data). The B channels can be used independently for voice or digital data, and the D channel is used for call control (replacing the POTS on-hook/off-hook, dialing, 20 Hz ringing, etc.) and low-speed customer data payloads. The two B channels can be linked to provide a 128 kbits/s channel. This combination of customer channels is often referred to as 2B + D.

The most common transport for ISDN is by a digital subscriber line (DSL) system that operates over a single ordinary two-wire nonloaded loop between the network and the customer premises. (DSLs have their own type of echo cancellers to enable four-wire digital to two-wire digital path conversions at the network and customer ends of the loop for the DSL signals.) The DSL extends four-wire transmission all the way to the customer. The customer terminal equipment can be many types of digital voice, data, or facsimile devices. An ISDN voice telephone set could be digital up to nearly the microphone and speaker. Side-tone coupling from microphone to speaker would be added to make the phone sound live just as in analog sets. Any echo that could get back to an electrical network echo canceller would be acoustic echo from handsfree terminal sets.

Because of the huge embedded base of analog voice telephone terminal equipment, the ISDN DSL can also attach to a terminal adaptor (TA) that connects to existing customer wiring, telephones, and analog terminals. This terminal adaptor contains the analog hybrid that would be at the central office for analog POTS. Electrical echoes can get back to a network echo canceller from the impedance mismatches between the terminal adaptor hybrid with its compromise balance network and the mixes of devices connected to the inside wiring. This is similar to wireless local loop and hybrid fiber coax connections at the customer premises.

There are recommendations for the on-hook and off-hook impedances of voiceband terminal devices, nominally 600 Ω off-hook at 1 kHz. When the number of devices off-hook can vary (even during a connection), the customer premises input impedance can vary considerably. A 6 dB impedance return loss is easy to get when connecting between the network termination, such as an ISDN terminal adaptor, and the inside wiring. For instance, three telephones off-hook yield a 200 Ω shunt loop impedance for a 6 dB ERL. This customer premises echo source is usually buffered by the two-way loss of the loop. Loops have an average of 4 dB loss at 1 kHz each way or 8 dB for a reflected signal. This 8 dB plus 6 dB mismatch yields an ERL of 14 dB and is usually of less concern than the office hybrid echo. The terminal adaptor may insert about 4 dB of analog (or digital) loss in both transmission directions to simulate the average loop loss to keep the end-to-end loss about the same as for analog service delivery. The 6 dB mismatch is then buffered when reflected back to the network as before. The impedance mismatch can be worse (less) than 6 dB at frequencies in the voiceband, especially at the low and high ends. If the voiceband attenuation in the adaptor is less or more than 4 dB, then the ERL seen at (by) a network echo canceller will change accordingly.

ISDN should normally use four-wire, full-rate 64 kbits/s µ-law or A-law encoded voice all the way to the customer premises. The speech quality delivered to the customer is toll quality. A DSL is allocated 0.5 ms for one-way transmission (or 1 ms two-way), so there is little additional delay over analog loop transmission. This very rarely puts an ISDN circuit beyond the 64 ms tail-end capacity of the TECO3264. In contrast, some other telephony arrangements that move the four-wire circuit path all the way to the customer premises use very low-bit-rate voice (LBRV) in the final link to the customer and thereby introduce the delay and nonlinearity that comes with the aggressive voice companding techniques. Examples of these are wireless local loop and wireless cellular systems, especially some of the new digital systems.

ISDN DSL technology has been ready since 1987. Some large business locations have been using ISDN for several years, particularly in Europe and Japan and in some large companies in the USA. Large scale residential and small business deployment just did not happen as forecasted. (1990 was proclaimed as the Year of ISDN in the U.S.A.). In 1998, ISDN was just beginning to be installed in significant numbers in some areas, since the price and the available services, such as Internet access, matched up with customer desires and willingness to pay.

Echo Canceller Applications (continued)

Internet Telephony Applications

There is a forecast that by year 2002 perhaps 25% of long distance telephony traffic will be packet voice over the Internet, rather than traditional circuit switched voice by long-haul networks. What began a short time ago as a low quality voice service offered for a low price by small, start-up telephony networks has quickly evolved to good quality voice service offered by traditional long distance network operators. The large long distance networks are starting to offer the Internet telephony services in competition with their normal circuit switched services in order to keep as much market share as possible.

For a price that is half or less than normal long distance services, customers were willing to accept less than the usual voice toll quality. Now after a few years of work on packet voice transport techniques, the voice quality delivered over an Internet phone call has greatly improved. At the same time the price has not increased. The reasons for the much lower cost have as much or more to do with political reasons than technical reasons. Internet telephony is not subject to most of the regulations and tariffs as the normal local and long distance networks and does not have the overhead costs of the embedded regulated networks. The current political climate seems to be that the U.S. government is not willing to increase control over the free Internet.

The first major voice over Internet protocol (VoIP) network was put into service in late 1998. Echo control was a major engineering concern.

Internet-to-Public Connection

In the voice telephony application, the Internet is taking the place of the circuit switched networks. From an echo control point of view, the concerns are basically the same when the Internet originated call connects to a caller on the regular public network. Transmission delay results from the distance, voice bandwidth companding, packet assembly, switching, and reassembly. Whatever the origin of the delay, it is still just delay like that from distance in traditional long-haul circuits. The result is that there needs to be echo control at the edge of the Internet connection facing the public network to deal with electrical echoes from hybrids. An echo canceller such as the TECO3264 with normal settings would fit just as it does at the edge of circuit switched networks. At the computer end, any echo would be acoustic echo from speaker energy being bounced back to the microphone. Acoustic echo cancellation functions built-in to the computer telephone board is the means for echo control at this end (see Packet Transmission Impairments on page 26).

Internet-to-Internet Telephony Connections

When an Internet call connects to another Internet caller using microphones and speakers connected to a computer, there may not be any hybrids to produce any electrical echo. Acoustic echo cancellation would be the sole means for echo control in this case.

Public Network to Public Network via Internet Long Haul

When the Internet is used as an alternative to the traditional long-haul circuit switched networks, the Internet telephony service provider must provide echo control just like the traditional network.

Packet Transmission Impairments

Packet transmission introduces its own new impairments to telephony. Like the usual digital transmission in traditional circuit switched telephony, individual bits or blocks of bits can be hit by interference, inverted, and be in error. In circuit switching, the customer-to-customer path is fixed for the duration of the call and the end-to-end delay is fixed. In packet transmission, each packet can in theory travel a different path through a complex, busy network between the source and destination with each experiencing different delays. Packets sent earlier may arrive after packets sent later. The packet reassembler must wait at least some minimum time, like the maximum expected delay difference, to put the signal back together in the original prepacket order. Also, whole packets may be lost or simply arrive too late to be placed back in the stream. For data signals where every bit may count, forward error correction techniques may be able to restore the errored data at the receiving end, else the data block may have to be retransmitted. In voice transmission, the resulting gaps from missing packets may be filled in with some least harmful default pattern, such as a quiet code. This would appear to an echo canceller as yet another source of noise and the packet delay would be just another source of delay.

Packet transmission is also used for permanent point-to-point connections in addition to being used for switching through a network. For example, most of the major cellular telephone system manufacturers use frame relay transmission with variable length HDLC-type frames between cell sites and base-station controllers. These packet paths carry customer voice and data payloads, call progress and control data, and cell site control and monitoring data. A problem that results is excessive delay (or latency in packet transmission terminology) for the voice transmission. The manufacturers are planning to migrate from frame relay to asynchronous transfer mode (ATM) with its short, fixed (53 bytes) frames to reduce the delay, starting in late 1998. Voice over ATM (VoATM) has been the subject of industry efforts in recent years to provide quality voice transport.

Another approach to reduce delay through packet networks is adding priority flags to the packets. Packets for delay sensitive services, like two-way voice, would be switched and transmitted before packets for delay insensitive services, like one-way broadcast video and audio, when there is contention for network resources. Planning for the next generation of Internet protocol (IP) includes priority marking and control.

ITU-T recommendation H.323 covers real-time audio, video, and data transport via packet based networks. H.323 requires G.711 (µ-law and A-law), G.722, G.728, G.729, MPEG-1 audio, and G.723.1 audio codecs (see reference 8).

Echo Canceller Background

Echo Canceller Terms

This section provides some background information for those new to voice frequency echo control in general and echo cancellation in particular. It discusses some of the terms and topics that occur very frequently in the technical specifications and literature. Some terms are merely defined with an explanatory sentence or two. Other topics are covered in more detail.

There are other good overall tutorials and primers on the subject of echo control, such as the ITU-T Recommendations G.168, G.165, G.164, and several others in the G series. This section is meant to supplement the tutorial material in these other echo cancellers or general telephony sources (see reference 2), not to replace them. Some good tutorials on various current telecommunication topics may also be found at http://www.webproforum.com.

The subtopics are intended to be read independently. As a result, a few points are repeated occasionally, but fuller discussions are under one heading.

General

The basic theory of echo cancellation dates back to 1966 when it was first proposed, to build a voice frequency echo canceller for one voice channel which resulted in a rack of circuitry that consumed a high amount of power. This was state of the art. Thirty years later, a single application-specific integrated circuit (ASIC) can handle 32 voice channels with milliwatts of power per channel.

An echo canceller forms a replica of the echo by passing the far-end signal through an adaptive filter that attempts to match the characteristics of the hybrid and the facilities between the canceller and the hybrid (see Figure 11, Echo Canceller Standard Four-Port Configuration below). This replica is then subtracted from the signal that enters the near-end port.

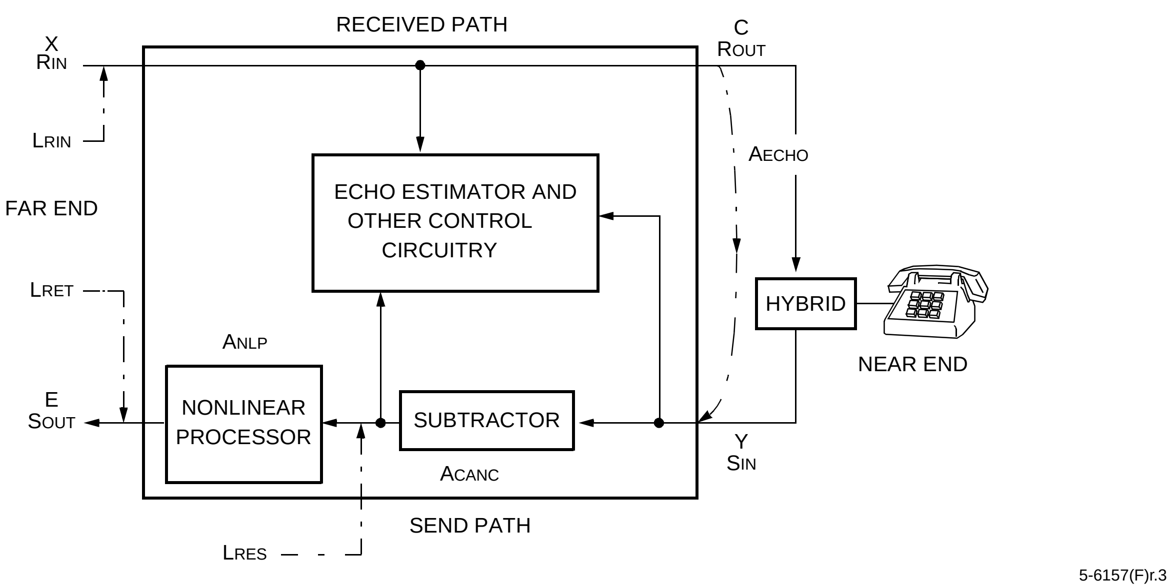


Figure 11. Echo Canceller Standard Four-Port Configuration

Figure 11 shows the reference model for echo cancellers that is consistent with ITU-T Recommendation G.168. In the definitions, L refers to the relative power level of a signal, expressed in dBm0, and A refers to the attenuation or loss of a signal path expressed in dB. The following terms are consistent with both ITU-T Recommendations G.165 and G.168.

Echo Canceller Background (continued)

Echo Canceller Terms (continued)

Convergence

Convergence is the process of developing a model of the echo path that will be used in the echo estimator to produce the estimate of the circuit echo by iteratively estimating the tail-end impulse response model in the H-register, calculating the estimated echo, measuring the residual echo error ($\hat{e}$/LRES), updating the impulse model, checking $\hat{e}$, etc.

Convergence Time

Convergence time for a defined echo path, is the interval between the instant a defined test signal is applied to the receive-in port of an echo canceller with the estimated echo path impulse response initially set zero, and the instant the residual returned echo level at the send-out port reaches a defined level.

An echo canceller should converge in a fraction of a second, with 0.5 s being relatively long.

Dispersion Time

Dispersion time is the time required to accommodate the band-limiting, and hybrid transit effects.

Echo Canceller

An echo canceller is a device used for reducing near-end echo present on the send path by subtracting an estimation of that echo from the near-end echo.

Echo Return Loss (ERL) (AECHO)

Echo return loss (ERL) (AECHO) is the attenuation of a signal from the receive-out port (ROUT/C) to the send-in port (SIN/Y) of an echo canceller, due to transmission level adjustments and transhybrid loss, i.e., the loss in the (near-end) echo path.

In voiceband echo canceller terminology, ERL is the signal level loss of a signal exiting the receive-out/C port and returning to the send-in/Y port. Since digital networks are usually set for 0 dB loss end to end at 1 kHz, the ERL value is basically the four-wire input port to four-wire output port loss of the hybrid network on a given call. Usually a single number, often 6 dB for a good network, is given and is understood to be the minimum ERL for any frequency within the voiceband for a given call. Return loss tends to be higher in midband, 500 Hz—2500 Hz, and lower at the band edges, below 500 Hz and above 2500 Hz, because of poorer impedance matching between the customer loop and the hybrid compromise balance network. For electrical reflections from a hybrid, the ERL is usually relatively stable in level versus frequency and in delay spread during a call, but may vary some. One example of the echo varying during a call is when the customer is very close to the central office, thus with a low loss loop. If a second telephone goes off-hook, the loop impedance seen at the hybrid will change, thus changing the ERL.

In non-echo-canceller telephone transmission terminology, the term echo return loss (ERL) is also used for a midband (500 Hz—2500 Hz) weighted return loss measurement referenced to the standard network-input/compromise-balance impedance (or other test impedance of choice). (The North American standard network two-wire input impedance is a 900 Ω resistor in series with a 2.16 µF capacitor, often referred to as 900 + 2.) This double definition of ERL sometimes results in minor confusion when the two transmission subcultures communicate. The corresponding low-band (200 Hz—500 Hz) and high band (2500 Hz—3000 Hz) impedance comparisons are called singing return loss low (SRL-LOW) and singing return loss high (SRL-HIGH). SRL-LOW and SRL-HIGH values tend to be lower than the impedance ERL measurement and these low and high frequencies tend to be the source of the echo canceller ERL value.

Echo Canceller Background (continued)

Echo Canceller Terms (continued)

Echo Path Delay (td, also Called Tail-End Delay or Tail Length)

Echo path delay (td, also called tail-end delay or tail length) is the delay from the ROUT/C port to the SIN/Y port due to the delays present in the echo path transmission facilities including dispersion time due to the network elements. In case of multiple paths, all delays and dispersion of any individual echo path are included.

Echo Return Loss Enhancement (ERLE) (ACANC)

Echo return loss enhancement (ERLE) (ACANC) is the attenuation of the echo signal as it passes through the send path of an echo canceller. This definition specifically excludes any nonlinear processing on the output of the canceller to provide for further attenuation.

ERLE is a measure of how much the echo canceller reduces the echo coming back into the send-in/Y port. Because of the nonlinear nature of the µ-law and A-law encoding/decoding process, the maximum ERLE that can be obtained is about 34 dB, which the TECO3264 can achieve. This 34 dB ERLE is still not good enough and a nonlinear processor is used to finish the echo control job.

Far End

Far end is the side of the telephone connection that does not contain the echo path on which the echo canceller is intended to operate. For long distance circuits, this is the long distance portion of the end-to-end circuit. The delay in the long distance portion may be hundreds of milliseconds.

H-Register

H-register is the register within the echo canceller which stores the estimated impulse response model of the echo path from the C/ROUT port back to the Y/SIN port.

Leak Time

Leak time is the interval between the instant a test signal is removed from the receive-in port of a fully converged echo canceller and the instant the echo path model in the echo canceller changes such that, when a test signal is reapplied to Rin with the convergence circuitry inhibited, the returned echo is at the defined level. Leak is necessary to bound the H-register response for very narrowband signals, such as single tones.

Near End

Near end is the side of the telephone connection which contains the echo path on which the echo canceller is intended to operate. Transmission facilities, switches, the hybrid, the loop, and terminating customer telephone set are included in the near-end side.

Echo Canceller Background (continued)

Echo Canceller Terms (continued)

Near-End Speech Threshold (NEST)

Near-end speech threshold (NEST) is the minimum attenuation of the signal between port ROUT and port SIN for the echo canceller to declare that only echo is present.

Nonlinear Processing Loss (ANLP)

Nonlinear processing loss (ANLP) is additional attenuation of a residual echo level by a nonlinear processor (NLP) places in the send path of an echo canceller. Because human hearing can detect speech-like signals that are very small, even in the presence of noise that may be higher in level, a nonlinear processor function is used to finish the job by adding more loss or cutting off the send path when only the far-end speaker is talking. The NLP can also insert comfort noise to maintain a live circuit sound. When the near-end speaker makes an utterance, even very brief, the nonlinear processor must quickly restore the send signal path. When near-end and far-end speech are both present (double talk), the ERLE should be sufficient for the near-end speech to hide the residual far-end echo.

Pure Delay (tr)

Pure delay (tr) is the delay from the ROUT to the SIN port due to the delays inherent in the (near-end) path transmission facilities. In this case, the delay associated with the hybrid is assumed to be zero.

Residual Echo Level (LREs) ($\hat{e}$)

Residual echo level (LRES) ($\hat{e}$) is the level of the echo signal that remains at the send-out port of an operating echo canceller after imperfect cancellation of the circuit echo. It is related to the receive-in signal LRIN by:

$L_{RES} = L_{RIN} - A_{ECHO} - A_{CANC}$

Any nonlinear processing is not included.

In the TECO3264, the residual echo is labeled $\hat{e}$.

Returned Echo Level (LRET)

Returned echo level (LRET) is the level of the signal at the send-out port of an operating echo canceller that will be returned to the talker. The attenuation of a nonlinear processor is included, if one is normally present. LRET is related to LRIN by:

$L_{RET} = L_{RIN} - (A_{ECHO} + A_{CANC} + A_{NLP})$

If nonlinear processing is not present, note that LRES = LRET.

Echo Canceller Background (continued)

Plain Old Telephone Service (POTS)

This is the telephone service that most people have at home. Most business voice telephone services are also a form of POTS. The main characteristics are: voiceband, telephones, dial tone, dual tone or rotary dialing, and 20 Hz ringing (other ringing frequencies are also used, especially outside of the United States). The vast majority of telephone services are POTS, carrying speech or, more and more, voiceband modem data or facsimile signals. Although various digital data services are growing rapidly, the embedded base of POTS is so huge that it will be the dominant service for many years to come. In the developing countries which have little or no telephony service, POTS type services will probably be the major service, but data communications will also be important.

POTS Loops, Metallic Pairs of Today

The two-wire metallic transmission facility that connects POTS (and other customer services) between the customer premises (not premise as it often incorrectly written) and the local telephone switching machine. Loops are usually constructed with twisted-pair copper wires. One main exception is the flat, copper-clad steel drop wire used for the last short link to the customer premises from a utility pole. The physical structure of North American loops varies considerably from one customer to another, varying from a few hundred feet to miles long. Fine gauge wire, 26 AWG, is usually used close to the central office. Coarser gauge, 24 AWG or 22 AWG is used farther out for longer loops. Thus, the typical loop has a mixture of gauges. Telephone cable is generally available in 500 foot reels. To reach the average customer on a 12000 foot loop the pair must be spliced two dozen times.

For flexibility in assigning pairs to customers as they apply for phone service and eventually move on, connected unused parallel shunts called bridged taps are used to connect a cable to more than one customer area. Added inductors called loading coils are installed on loops longer than 18000 feet to reduce transmission loss. Longer loops in the U.S.A. are gradually being replaced with digital loop carrier (DLC) transmission/multiplexer systems that bring the four-wire hybrid interface out from the local switch to within 12000 feet (about 4 km) or less of the customer. European loops tend be shorter, and to use coarser gauge wire with no bridged taps or loading coils.

Sometimes old technology is around for a long time. As recently as the mid-1990s, and possibly still in 1998, remote villages in the South American mountains are linked to civilization by the one-iron-wire-with-earth-ground-return telephone transmission loop used in 1878. These circuits extend from the closest town with some modern telephone service up into the mountains for many miles and may link more than one village in a party line. (Such systems were still in use in the U.S.A. in the 1930s.) Reports are that the noise levels on these lines are very low and voice transmission is clear and quiet, probably because there is no source of electrical interference from any other electric systems of any kind. These ancient systems are good candidates for replacement by solar powered earth station sets for low-earth orbit satellite systems, thus skipping 120 years of intermediate technology. See Echo Cancellers in Nontraditional Networks, Connections to PSTN on page 19 and Delay on page 35.

Line

A telephone transmission link that has customer terminal equipment connected to at least one end. A POTS loop is a line. If both ends of line are connected to customer equipment and has no switch connections, it is a private line service, or basically an intercom. Lines tend to be at least partly two-wire, from the central office out to the customer premises. Any interoffice links of private lines constructed today will use digital four-wire facilities thus introducing two-wire to four-wire hybrids.

Trunk

A telephone transmission facility connecting two switches. The trunk is made up of the terminal electronic equipment (amplifiers, multiplexers, transmitters, receivers, etc.), the media (wire, radio, optical fiber), and possibly intermediate amplifiers. The switches may be local, long distance, or private network switches.

Modern trunks are almost always four-wire. Trunks connect the local network to the long-haul networks and to private switched networks.

Echo Canceller Background (continued)

Facility

Refers to a telephone transmission link (line, trunk) including the terminal electronic equipment and the interconnecting media (wire, optical fiber, radio). Some facilities also have intermediate electronics when the distance is too great to span between the terminal equipment. In the past the signal delay through a facility was due mainly to the propagation time through the media with a small contribution due to the attached electronics.

Switch

The machine that connects one telephone line or trunk to another line or trunk to build up an end-to-end call. Modern digital switches are basically very specialized, ultra-reliable computers that connect four-wire, bidirectional 64 kbits/s DS0 paths to one another. Analog, two-wire switches are still in service and still work very well for local POTS. The local switch is usually called the central office and connects POTS lines to other local POTS lines or to long distance/toll trunks. The long distance or toll switching machines usually just make trunk to trunk connections among themselves and other toll switches or local switches. Some toll switches may make direct connections to large company customer private networks and bypass the local switching systems. Network echo cancellers are usually located on the long distance side of long distance/toll switches pointing with the near end or C and Y ports facing the local switch to cancel the local network hybrid echoes. The far end or X and E ports face back toward the toll network.

Signal delay through an old analog circuit switch was very small and was mainly due to propagation through the intra-building wiring. Signal delay through a packet based switch, such as ATM, can be considerable.

Central Office

Strictly speaking, the local telephone switching machine, identified by the first three digits (called NNX) of the local telephone number in the U.S.A. as in the 879 of 879-6143. Central offices are housed in buildings called wire centers because the local network of customer loops radiate outward from there. A wire center may contain one or several central office codes with up to nearly 10000 customer lines each, served by one or more switching machines. The term central office is often used to refer to the whole building as well.

Public Switched Telephone Network (PSTN)

PSTN is the name for the collection of telephone networks (local and toll switches and trunks) that make up the public switching network, originally intended for voice service. Many large companies, governments, the military and other entities have their own private telephone and data networks that carry most of the traffic between their locations. For interconnections outside their networks, they may have one or more gateway connections to the PSTN. Hopefully, these private networks use the same care in echo control as the PSTN and long distance networks. Some military networks use four-wire switches, facilities, and station sets which enables them to avoid echo control problems.

PSTN Circuit Switching versus Data Network Packet Switching

The type of switching used in the PSTN is called circuit switching which means that when a customer places a call a path is constructed from the customer’s switch port connection to the called party switch port and it is dedicated to that call as long as both parties are connected, even when no one is talking. The path through the network may vary greatly from one call to the next for the same pair of customers, but it is fixed for a given call. In contrast, data networks have been designed to use packet switching where the connection may vary from one data packet to the next during one session between two customer data ports. More and more voice traffic is being carried on packet oriented data networks, with the Internet being the best known example. The delay may vary between data packets carrying voice information.

Echo Canceller Background (continued)

Echoes

It can be exciting and fun to hear one's voice come back as an echo from a canyon wall, a cliff, or big building. However, to hear an echo during a telephone call can vary from merely distracting, to annoying, to disrupting the flow of conversation completely.

There are three conditions necessary to have an echo perceived: (1) some sort of reflector which redirects some of the passing energy (electrical or sound) back toward the source, (2) sufficient returned energy relative to speech and background noise levels to be detected and (3) sufficient round trip time delay between the utterance and its return to the listener's ear. How bad a telephone call echo is perceived depends on the relative loudness of the echo and also on the time delay.

For telephone echoes, the reflector is most often a device called a hybrid which converts from a four-wire electrical transmission path to a two-wire transmission path. (See Four-Wire Circuit on page 37 and Two-Wire Circuit on page 37.) Practically every telephone customer is connected to the local serving telephone company switching machine by a single pair of metallic (usually copper) wires, or a two-wire path. This pair of wires is a bidirectional path carrying both directions of signal at once. Modern telephone switching systems and transmission systems use two unidirectional paths internally called a four-wire path. These four-wire paths may be on copper wires, fiber optic cables, or radio waves. (Some military telephone networks use all four-wire circuits and four-wire telephone sets and thus need no hybrids.) So when a customer call connection is made, there must be a hybrid to get from the first wire pair into the telephone network and another hybrid at the far end to connect to the second customer wire pair. In modern networks, the path through the network or networks is usually a four-wire path all the way between hybrids, even half-way around the world, or farther if a satellite link is involved.

As mentioned elsewhere, the four-wire to two-wire conversion process in hybrids is not perfect. Some of the energy that is supposed to go outward from the incoming four-wire side to the two-wire side leaks across and goes back out on the outgoing four-wire port. This trans-hybrid leakage, called echo return loss, in echo control termin ology, is usually very small relative to the outgoing two-wire port signal. Even a very poor, nearly completely mis-matched hybrid would have a 6 dB echo return loss, meaning the reflection is down by a factor of four. For reason-ably well designed systems, the echo return loss would run closer to 12 dB, or a factor of 16 reflection. However, the human ear-brain signal processing system is very, very sensitive when it comes to detecting signals that sound like they might be speech. A reflected signal that is 40 dB or even 50 dB down by the time it returns to the speaker can be heard. So far even for a network that has very good hybrid performance may need echo control if the round trip delays are large enough.

When is a reflection perceived as an echo? If a sufficiently large echo comes back to the speaker’s ear in less than 10 ms, it is not recognized as an echo. (Bats and dolphins must be able to distinguish much finer time differences in sound echoes than 10 ms in order to locate bugs and fish.) It merely blends in as part of the side-tone designed into telephone sets. (Sufficiently large can be 40 dB down. If a signal in a bad network comes back larger than the original, there are different problems including the possibility of singing like a bad public address system.) As the delay increases toward 30 ms, the circuit starts to have a hollow sound. Nearly everyone will hear a reflection as a a separate sound with 50 ms delay if the reflected signal level is not much lower than the background noise. (50 ms round trip delay for a sound echo means the reflecting surface is about 200 feet away.)

If an echo returns with 0 dB attenuation inside a two-wire-to-four-wire-to-two-wire circuit, the circuit can become an oscillator or start singing. This is a familiar effect in audio public address systems that squeal until the gain is turned down. Usually, there is enough ERL that singing doesn’t happen often in telephone networks.

Telephone connections which have the possibility of having a round-trip delay of anything like 30 ms or more must have some form of echo control to deliver a circuit over which customers can carry on a satisfactory conversation. Before the echo canceller was invented by Bell Labs and the development of very large scale integration made complex devices such as the Lucent TECO3264 practical, several approaches to echo control were tried with vary-ing degrees of partial success. Getting rid of the reflection with perfect hybrids is impossible. Providing very high quality hybrids for every one of the hundreds of millions of customers is impractical. The past approaches usually involved adding full time loss to the long distance trunk links that increased with distance. The North American echo control plan was called via net loss (VNL). These long distance trunks are not dedicated to any one customer, but are shared among all calling customers on demand under control of the switching systems in the local and long

Echo Canceller Background (continued)

Echoes (continued)

distance networks. Echo suppressors were developed that attempted to detect when the far-end party was talking but not the near-end party and switch in a fixed loss. When the near-end party started talking the suppressor tried to switch out the loss without the humans detecting it. It was better than nothing and the only practical solution at the time.

Being able to share the complex, expensive echo suppressors by placing them back in the edges of the long distance portions of the networks made the per customer, per call cost feasible. (For small [small in numbers of circuits, not in physical extension or delay aspects] networks, where there is not the ability to share relatively complex echo control systems, precision hybrids are used as the echo control devices of choice.)

With modern echo cancellers, it is possible to design a long distance network that has a 0 dB loss from end to end and insert echo control at the ends that interface with the local networks having the troublesome hybrids. An echo canceller in a traditional long distance network will be placed at each end of the link with each facing the nearest local network. The round trip delay between the echo canceller position and the local network hybrids is seldom more than 64 ms. Each echo canceller only has to account for this local tail-end delay to its hybrid, not for the delay for the portion of connection between the echo cancellers. The long distance portion can be hundreds of milliseconds if satellite links or very heavy signal processing or bit-rate compression is involved. If the total delay gets into large fractions of a second, even with echo cancellers, the human speaking protocol may have to become more formal and less natural in flow with military type conversations with over at the end of each sentence so the listening party knows when the speaker is done.

For each call, the long distance trunk echo cancellers will be connected to a different local telephone trunk and then to a different local customer with his own hybrid, two-wire loop and customer premises equipment. Thus, for each call the echo cancellers may see a different hybrid with different echo return loss level with a different delay from the previous calls. The echo cancellers must adapt as quickly as possible in order to cancel echoes at their respective ends, hopefully before the human ear hears any echo. The hybrid echo characteristics do not tend to change much during a call, but when changes do occur, the echo canceller must recognize them and adjust as necessary to maintain quality.

When one person speaks and a second party listens until the first party is finished speaking, then the second party speaks while the first party listens, etc. This is called single-talking. People do not talk in this manner. Very often, both parties speak at the same time. It may be as simple as a yes or no or just a “hmmm” to indicate that the listener is indeed still listening. If the conversation becomes more excited, both parties may be speaking at the same time. These situations are called double-talking. If humans only single-talked, the task of designing an echo control device such as an echo canceller, would be much simpler. The TECO3264 employs advanced detection techniques and algorithms to reliably determine when the far end speaker is speaking or not, when the near end speaker is speaking or not, so that the echo canceller can be trained at the appropriate times, the nonlinear processor can be activated/deactivated, and comfort noise inserted and removed at the right times. Low ERL values make it harder to distinguish between an echo and a quiet speaker at the near end. The whole double-talking accommodation process by the echo canceller should be transparent to the listening/talking customers.

Telephone call connections in the modern world are usually made up from several pieces supplied by several telephone network providers, starting with the callings customer’s local operating company, by a long distance network provider and finally by the called customer’s local operating company. The long distance portion of the call may be made up from one or more national links, international links and links by provided by other circuit providing entities. Wireless mobile calls are linked into the landline networks and then to a landline customer or back out to another wireless customer.

(In over-the-air television transmission, echoes from hills, buildings, airplanes, etc., are seen as visual ghosts on the television screen. Ghosts can occur in poorly engineered cable TV systems as well from reflections off poor cable connections, or improperly terminated cables.)

Echo Canceller Background (continued)

Delay

Formerly, the transmission delay between the time a signal left the transmitter and arrived at a receiver was just a result of the fact that physical things can only travel at finite speeds. Radio signals and light in air travel at nearly the speed of light in a vacuum (300000 km/s or 186000 miles/s in round numbers), the maximum possible speed. Milliseconds are the common time measure in echo control. In 1 ms, a radio wave travels 300 km/186 miles and a sound wave travels about one foot. An electrical signal in plastic-insulated, twisted-pair copper wire travels about half light speed, or about 150 km or 100 miles in 1 ms. For echo purposes, the round-trip delay out and back is what counts, so the round trip distances for 1 ms propagation delay would be half the one way distances. So an electrical signal traveling 75 km or 50 miles trip through wire pairs would experience a round trip delay of about 1 ms if it were reflected back toward the source from the far end.

The delay in the average North American two-mile loop, or even two loops connected together for a call, does not come close to causing an echo problem for local calls with only microseconds of delay. Loops in Europe tend to be shorter and would be of even less concern for delay. As an extreme example, links through a geosynchronous satellite at 22300 miles/35900 km orbital altitude above the equator results in a 0.25 s round trip delay which greatly affects two-way speech communications. This large satellite delay does not bother one-way communications services, such as audio and video broadcast, data transmission, or paging. (Data transmission protocols do have to account for the delay.)

In 1998, there were low-earth orbit (LEO) voice communications satellite systems nearly ready for service. These satellite orbits tend to be about 1400 km/870 mile altitude, which results in a round trip radio propagation delay of about 9 ms just from the air/space path. The slant range from the earth-based transceivers to the satellites will usually be longer than the satellite altitude since they will not be directly underneath. So the delays will usually be longer than the minimum 9 ms. Some systems plan intersatellite relays of signals by radio or laser before returning to earth which will result in still longer propagation delays. These satellite systems tend to use low-bit-rate-voice speech compression to preserve analog and digital bandwidth. Use of digital forward error correction will add more still more delay. Echo control is a consideration in the LEO satellite voice systems.

Delay through the terminal equipment and switches connected to the copper, fiber, or air/space paths used to be negligible, being typically a few ms. For the high-bit-rate digital subscriber line (HDSL) there is an allocation for 0.5 ms of one-way delay due to the line encoding and decoding. This was considered a large delay for terminal equipment in 1991. This is not the case any more with tens of ms being introduced by digital speech compression, packet and ATM transport, and digital forward error correction applications. The asymmetric digital subscriber line (ADSL) has an allocation of 20 ms for one way delay when the error correction option is used.

Echo Canceller Background (continued)

The Telephone System: Not Perfect for Very Good Reasons

Today's telephone system goes all the way around the world and from the northernmost to the southernmost inhabited points on the earth. It involves an incredible investment in money, time, material and human ingenuity. The system is not perfect, but it is ideal in the sense that it provides acceptable voice service for a price the customer is willing to pay. Devices like echo cancellers exist because the system is not perfect.

The design and implementation of the telephone system, like any engineered system or tool, involves choices and balances and trade-offs among many conflicting concerns like cost, performance, state of the art, timely deployment, reliability among others. The design choices have a history starting with the pioneers like Bell, Grey, and Edison over 100 years ago.

In the beginning, even a barely perceivable hello heard across town was a miracle. (The word hello was coined for telephone use, Bell wanted to use ahoy.) As the technology improved, that quickly became not good enough. To start with the transmission facilities were one iron wire with an earth ground return (still used and still works well in some South American mountains in the 1990s). Then Bell came up with insulated, twisted copper wire pairs. Edison, Grey, and others developed better microphones, earphones, and other equipment. Stowger, a funeral home director, invented the electromechanical switch controlled by customer dial pulses. The quality of voice transmission improved and the transmission distance increased and increased until today.

At each stage design decisions had to be made. For the telephone to be successful, it couldn't be just a business or rich man's tool. It had to be cheap enough for everyone to be able to afford it. Universal service was a real goal. The vast bulk of the equipment and investment was and is in the local network serving local POTS customers. While four wires to the customer and four wire terminal equipment (as used in some military systems) would be nice, two-wire would serve and be a lot cheaper. Local switches were also two-wire. While short-toll trunks could be two-wire, long distance transmission required four-wire transmission and switching machines so amplifiers could be inserted. This meant hybrids had to be introduced into the interface between long distance and local networks.

Hybrid design, impedance matching, and return loss control were developed into fine arts with exacting design rules to minimize echoes on long distance calls where the distance meant transmission delays. Hybrids in two-wire toll switches had to meet terminal balance requirements where the average midband return loss for a group of trunks had to be 18 dB. The minimum for any one trunk had to be at least 13 dB translating to a transhybrid loss of at least 21 dB. Since the long distance network investment is considerably smaller than the local plant, this was the logical place to make up for necessary compromises in the local systems.

As local four-wire digital switching machines began to replace the two-wire analog switches in the 1970s, the four-wire to two-wire hybrid interface moved out the interface to each local customer's loop at the central office end or with digital loop carriers out to a location in the field closer to the customer premises. These local customer interfacehybrids do not, indeed can not, get the intense degree of impedance matching and return loss control that the far fewer toll hybrids back in the network once received.

One aspect of telephone system design that was not compromised is reliability. Telephone equipment is expected to be installed, turned on, and work night and day for decades. The failure of a telephone switching system for any reason other than a natural disaster is a news event and the failure is analyzed in detail to help avoid similar failures. The most common large failure in telephone systems occurs when a construction company digs up a telephone cable by accident. Compare this to normal computer systems where a day without a crash or data loss of some sort is a success. The TECO3264 is designed for mean time between failures measured in decades.

Some day the voice band telephony network may have audio compact disk quality (16-bit linear, analog bandwidth of 20 Hz to 20 kHz) and all four wire transmission with no reflecting hybrids, but not for a long time. In the meantime, devices such as echo cancellers make good engineering and economic sense.

Echo Canceller Background (continued)

Two-Wire Circuit

Two-wire in telephony literally refers to two metallic wires in a pair that are used for a two-way communication circuit. The wires are usually copper and usually twisted to reduce outside electrical interference to and from other pairs and other electrical noise sources. A two-wire telephony circuit is usually bidirectional and full-duplex; that is, signals travel in both directions (in/out, left/right, east/west) at the same time. Voice circuits are usually symmetric in bandwidth, that is they have about 3000 kHz or 64 kbits/s available in both directions to carry signals. (Data channels are not necessarily symmetric. A 15:1 asymmetry in two-way data bandwidth requirements can happen.) Speech and voiceband computer data modem signals are two common signals.

Nearly all of the several hundred million telephone customers in the world are connected by a single two-wire pair from their home or business all or at least part of the way to their local telephone switching system. This two-wire connection is often called a loop, referring to the full loop path for network supplied direct current to power the station sets and to provide loop closures to indicate the handset has gone off- (or on-) hook.

Large business or government entities may lease four-wire wire, wireless, or optical facilities from telephone service providers and bypass the local telephone company and its two-wire loop plant.

Four-Wire Circuit

Four-wire in telephony used to always mean two pairs of wires used for a two-way communication circuit. Signals on each pair are unidirectional; that is, one pair carries the signal in one direction (in to out, left to right, east to west) and the other pair carries the signal in the other direction (out to in, right to left, west to east). The telephone system has evolved from an all analog network to nearly all digital except for the customer two-wire loop and home telephone sets. The term four-wire now usually refers to the two unidirectional paths inside a digital transmission system such as DS1 or E1 based multiplexed carrier systems or inside digital switching systems. Voiceband speech or data signals are converted to binary pulse coded modulated signals usually (but not always) at a 64 kbits/s DS0 rate. The signal bits appear between logic gate outputs and ground, not on two-wire pairs, but the term four-wire circuit is still used to indicate two unidirectional paths.

Telephone Station Sets, Hybrids, Side Tone

Telephone sets are both two-wire and four-wire. The connection to the telephone network for the vast majority of analog station sets is via a two-wire metallic cable pair called a loop (see Two-Wire Circuit above). The connection to the human side is four-wire, with two wires to the microphone and two-wires to the earphone. Thus, the ordinary telephone must have a two-wire to four-wire hybrid conversion circuit inside it. If the station set hybrid were perfectly matched to the impedance of the connecting two-wire loop, none of the speaker's voice energy from the microphone would go to the earphone. In fact the hybrid has a deliberate, carefully controlled mismatch to the loop, causing a fraction of the speaker's voice energy to be heard in the earphone. This designed-in microphone-to-earphone reflection is called side tone. Since the delay is very, very small, the speaker/listener is usually not aware that it even exists, unless the circuit fails and there is no side tone in which case the phone sounds dead. Telephone systems designers discovered that if the customer hears his own voice in the earphone, he will tend to control his speaking volume much more so than if he doesn't hear his voice. This more controlled range of speech volumes helps in designing the telephone equipment and setting the network operating parameters. Side tone helps the customers feel better and the network work better and cheaper.

Echo Canceller Background (continued)

Hybrids: General Description

Hybrid, or more completely two-wire to four-wire hybrid, or two-wire to four-wire termination set, as used in voice-band telephony, is a balanced bridge network used to convert between two-wire telephony circuits and four-wire circuits. Hybrids are used at the two-wire ports of digital switching machines and digital carrier systems which are designed as four-wire systems internally.

Signals coming in from the customer to the hybrid to its two-wire port are sent out on the outgoing unidirectional port on the four-wire side. Signals coming into the hybrid on the incoming four-wire port of the hybrid are directed to the two-wire port to go out toward the customer. Ideally none of the incoming four-wire signal gets across to the outgoing four-wire port. However, hybrids are real, physical devices and some of the incoming four-wire signal does leak across to the outgoing port. If the leak, or reflection in echo terms, is large enough, the other customer at the far end can hear his own voice, even 50 dB down. The loss from incoming four-wire port to the outgoing four-wire port is called, echo return loss (ERL). In non-echo-canceller telephone transmission terminology this leak/loss is called transhybrid loss. The ERL as measured at an echo canceller includes any signal level adjustments, up or down, in the network going to the hybrid or returning as an echo, in addition to the transhybrid loss.

If the round trip delay is greater than about 10 ms and less than about 30 ms, the reflected signal is perceived by the far-end listener as a hollow sound. Delays greater than about 50 ms result in the reflected signal being heard as a separate echo from the original utterance. Hybrids are dominant, but not the only source of electrical echo in telephone circuits.

It is possible to design and build hybrids that increase the incoming port to outgoing port loss, or echo return loss, to >30 dB. Such hybrids are complex and expensive and are not justified by the vast majority of calls that are local, short haul with low delay and no perceivable echo. Such hybrids are used for special customer circuits.

Not shown on most simplified block diagrams is a second two-wire hybrid port to which is attached a balance network that attempts to match/balance the input impedance of the two-wire cable on the signal port. Impedance is the ratio of the signal voltage to the signal current on the wire pair versus frequency. For any one wire pair, the impedance varies drastically from 300 Hz to 3000 Hz (see Voiceband on page 41). The impedance also varies with other variables such as the temperature of the cable, the number and type of telephone sets and other devices connected and active at the customer end on a given call.

The physical structure of loops varies considerably from one customer to another, varying from a few hundred feet to miles long, construction of fine gauge wire (like 26 AWG), to mixtures of gauges, connected unused parallel shunts called bridged taps, to added inductors called loading coils on long loops. The input impedance among loops varies considerably. The approach to a reasonable hybrid balance network that does a good enough job on local calls, is an electrical circuit called a compromise balance network that provides a hybrid mid-band (500 Hz—2500 Hz) echo return loss of about 15 dB averaged across the population of loops connected to a switch and with a minimum in the range of 6 dB for the worst loops. The return loss or impedance balance degrades below 500 Hz and above 2500 Hz. The minimum return loss, not the maximum or average, is the key because that is the cause of the trouble. Ninety-nine great hybrids and one bad is still bad. A hybrid with 50 dB echo return loss across almost all the voice band, but only 2 dB at any one frequency is still bad.

More modern digital switches segregate the loops into two groups, nonloaded (about 80% in the U.S., the shorter loops) and loaded (about 20%, the longer loops, >18000 ft, with inductive loading). (Europe has no loaded cable.) Using separate nonloaded and loaded compromise hybrid balance networks provided about another two dB of echo return loss on the average and some improvement in the minimum for a given loop plant served by a switching machine. This is still not good enough to prevent echoes when the hybrid and the connecting two-wire loop are part of a long distance circuit. However, since the switch has uses two hybrids to connect two customers, the total increase of 4 dB is critical to the design of a 0 dB loss from switch two-wire output to two-wire output. Otherwise, the switch would have to insert loss or often sound like a barrel.

Echo Canceller Background (continued)

Hybrids: General Description (continued)

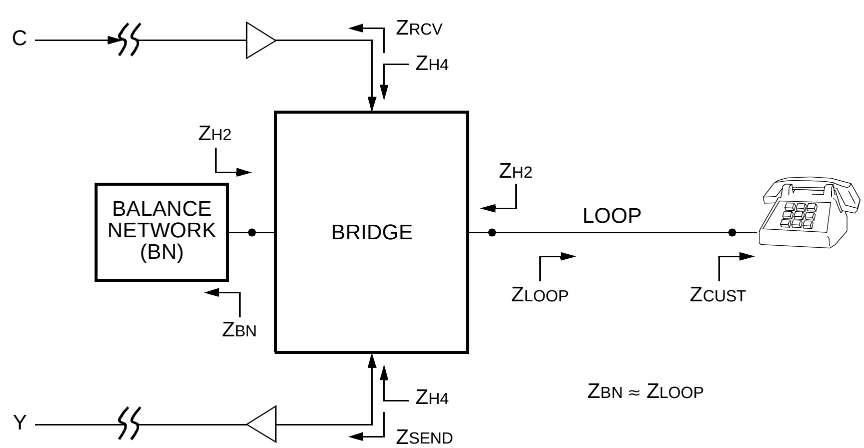
The trend in hybrid echo return loss seems to be deteriorating rather than improving with time. The traditional long distance networks depended on the careful, reasonably predictable hybrid echo return loss as described above and provided by local switching machines. With the introduction of voice telephony by packet systems, over the internet (the connection to other private and public networks in which the concept of good overall network and hybrid design for echo control is a lost art) there have been instances reported of echo return gain! Fortunately, a modern echo canceller, such as the TECO3264, has parameters that can be adjusted to accommodate less than desired network performance.

Hybrids: Electrical Description

As mentioned in the general discussion of two-wire to four-wire hybrids, they are impedance balance networks. Figure 12 below depicts a hybrid network as a block with the impedances with it interfaces and the impedances that it presents to the outside.

Point C is the receive-out/C port of an echo canceller system. The signal travels through the intermediate network transmission and switching equipment (which is not shown) to the four-wire input port of the hybrid. Likewise, point Y is the send-in/Y port of an echo canceller that receives the signal from the four-wire output of the hybrid via the intermediate (unshown) equipment.

One two-wire output port of the hybrid connects to the metallic loop that in turn goes out from the central office and connects to the customer's inside wiring and telephone sets. The second two-wire port of the hybrid connects to an impedance balance network (BN). This balance network is usually not shown in telephone connections. The task of the balance network with its input impedance, Z_{BN}, is to match as closely as necessary the input impedance of the loop as seen at the hybrid. This loop impedance, Z_{LOOP}, is a complex function of frequency. It varies from one customer loop to the next. It can vary some with the outside temperature. It can vary with the number of telephones being used on a given call that change the input impedance to the customer premises, Z_{CUST}, especially for short, low loss loops. As discussed above, Z_{BN}, does not have to be a perfect match to Z_{LOOP} for local telephone connections. For older, analog two-wire switching systems in North America, a single compromise balance impedance, 900 Ω + 2.16 μF (see above). Modern local digital switches segregate the loops and hybrids into nonloaded and loaded populations and use separate nonloaded and loaded balance networks. These two new balance networks are still simple with only three or four passive resistors and capacitors.



5-7082(F)r.2

Figure 12. Hybrid: Two-to-Four Wire Converter

Echo Canceller Background (continued)

Hybrids: Electrical Description (continued)

The two-wire ports of the hybrid are usually designed to present the standard 900 Ω + 2.16 µF impedance, Z_{H2} on both sides. This can be shown to be the transformation of the input impedances, Z_{NET} (typically 600 Ω) in parallel, looking back into the network from the four-wire ports of the hybrid plus some additional resistances and capacitance. The input impedances of the four-wire ports of the hybrid, Z_{H4}, are the parallel, transformed versions of Z_{BN} and Z_{LOOP}. The transformer windings are connected such that the phases of the loop and balance network port reflections tend to cancel at the outgoing four-wire port.

Skipping the details, it can be shown that ideally, half the power input to the four-wire port goes to the loop, half to the balance network and none to the send port. Likewise, half the signal power coming in from the loop goes to the receive four-wire port of the hybrid where it is dissipated in the output of the connecting amplifier. Half the power from the loop goes out the send side of the hybrid as desired. None of the loop input power ends up in the balance network.

Skipping the details, the four-wire input to four-wire signal transmission function of the hybrid turns out to be of the form, $(Z_{BN} + Z_{LOOP})/(Z_{BN} – Z_{LOOP})$. This is the same form as the return loss (RL) or the inverse of the reflection coefficient between a transmission line of impedance Z_{LOOP} and a termination of Z_{BN}. The other impedances around the hybrid drop out of the THL, if they are as shown. Ignoring the phase of the complex function, the transhybrid loss (THL), or echo return loss, has the form (K + 20 log(magnitude(($Z_{BN} + Z_{LOOP}$)/($Z_{BN} – Z_{LOOP}$))). For older, two-transformer hybrids, K was usually about 8 dB when miscellaneous losses were included. THL would then be about 8 + RL dB as a rule of thumb.

For the North American loop plant, the 900 + 2 balance network has an average midband return loss of 11 dB and a minimum of about 6 dB. The high- and low-band return loss values are about 3 dB less, or 8 dB average and 3 dB minimum, which yields a minimum THL of about 8 + 3 = 11 dB. Note that this is completely adequate for local telephony, but not for echo control purposes when the local link is part of a long haul circuit. The segregated non-loaded and loaded balance networks improve this by about 2 dB to yield a minimum THL or ERL of about 13 dB, which is still not nearly good enough for echo control.

Note that a completely mismatched hybrid with a short or open balance network, would have an transhybrid loss of 8 dB. A lossless hybrid would have a minimum transhybrid loss of 6 dB, since half the power (3 dB) is first split between the loop and the balance network. The reflected power differences from the balance network and the loop get split in half, another 3 dB, again when going out both four-wire ports. Miscellaneous losses tend to add another 1.5 dB to 2 dB for the rule-of-thumb constant 8 dB, mentioned above.

For old-style, two-wire transformer hybrids, there is also a fifth, low-frequency port for feeding dc current from the office 48 Vdc battery supply to power the telephone down the loop and for 20 Hz ringing current to ring the bell. These power and signaling currents are fed through large inductors and across shunt capacitors that connect the center tap transformer windings. This hides the impedance loading effects on the hybrid at voiceband.

Two-wire transformer hybrids are out-of-style now because they tend to be rather large, with plenty of iron in the cores and copper wire in the windings to avoid magnetic saturation of the core from the dc and ringing currents fed out through the hybrid. Small one-transformer, electronic hybrids with external feed for dc and ringing currents are now the norm. The transmission analysis for transhybrid-loss/echo-return-loss is still the same. The old, reliable, rule of thumb of an 8 dB minimum transhybrid loss seems to be gone, with the number these days often coming in at 6 dB, 3 dB, or less.

Echo Canceller Background (continued)

Voiceband

Voiceband in telephony usually means the signal spectrum from about 300 Hz to about 3000 Hz, sometimes given as 3400 Hz. Human speech in the air may go from approximately 20 Hz to 10 kHz. Two-wire copper pairs can carry electrical signals from 0 Hz/DC to 30 MHz and more. For human understanding of speech, the range from 300 Hz to 3000 Hz is adequate for acceptable quality. The connecting circuits in the telephone sets, the transmission systems and switches filter out frequencies below 300 Hz and above 3000 Hz to reduce the overall cost and complexity of the network. The sampling that is done as part of the analog to digital conversion cuts off rapidly above 3000 Hz and the signal is essentially gone by 3400 Hz. Legal tariffs that specify voiceband telephony services are often written in terms of the loss at 404 Hz and 2804 Hz relative to the end-to-end loss at 1004 Hz. For example, a high quality private line service could allow up to 1 dB above and 4 dB below 1004 Hz levels at the 404 Hz and 2804 Hz points measured end-to-end. At least one telephone network does some boosting of the signal below 400 Hz and above 3000 Hz to restore more natural sounding speech. This boosting can aggravate echo control problems because hybrid balance tends to degrade at the low end (below 500 Hz) and at the high end (above 2500 Hz) of the voiceband.

POTS Call Control Signaling Transport and Echo Cancelling

In addition to the customer signals (voice or data carried within the DS0 channel), the network has to transport additional information to allow network control of the call setup, alerting of the called customer, and final disconnect. For POTS calls, such information includes on-hook and off-hook signals to indicate whether the customer has picked up the telephone handset to initiate a call or to answer an incoming call, and the presence of ringing, typically pulses of 20 Hz, coming from the originating switch. For old-style rotary telephone sets using dial pulse addressing for sending the calling number to the network, the on-hook/off-hook signal is modulated, typically at 10 pulses per second. Modern systems tend to use dual-tone multifrequency (DTMF) dialing within the voiceband for sending phone number addressing. When the network is connecting to a private customer switch, called a private branch exchange (PBX), additional call control signaling functions are required.

The on-hook/off-hook and ringing signals are at very low frequencies, from 0 Hz/dc to tens of Hz including the harmonics of the dial pulses and 20 Hz ringing pulses, and pulse rise and fall times. These signals are very much lower in frequency than the typical lower cutoff frequency for the voiceband, which is about 300 Hz. The on-hook/off-hook and ringing are out-of-band relative to the voiceband while DTMF signaling is in-band. For digital loop carrier systems connecting a customer premises to the telephone network, this out-of-band call control signaling must be transported back and forth along with the customer voiceband payloads by some means.

DS1 systems and E1 systems tend to use very different techniques to transport call control signaling. DS1 systems rob bits away from the 64 kbits/s voice signal to carry control signaling in a robbed bit signaling scheme. E1 systems gather up the control signaling from 31 of the DS0 channels, and multiplex it all together in the 30 second common signaling channel.

DS1 systems take the least significant bit from the eight voice band bits from every sixth 125 µs frame. This provides 64/(8 x 6) kbits/s or 4/3 kbits/s of bandwidth for call control information in each direction, which is plenty for call control purposes. Older DS1 systems divided this into two channels, A and B. Newer DS1 systems, using extended superframe (ESF), divide the robbed bits into four channels, A, B, C, and D, in each direction to allow for more complex interactions between switching systems and other control functions.

Note: The 20 Hz ringing pulses are not sent by PCM encoding, but just the presence/absence of the 20 Hz bursts are detected so that they can be recreated at the far end.

This DS1 robbed-bit signaling approach leaves a net voiceband capacity of (64 – 4/3) kbits/s while introducing a little low level noise. With the robustness of the µ-law encoding, the human listener never hears anything wrong.

What happens if an echo canceller and a carrier transporting call control signals were to encounter? For E1-based systems, it is simple. With the TECO3264, the common signaling control channel can be set to 64 clear/bypass mode and the signaling passes through the echo canceller untouched. If a DS1-based system must preserve the robbed-bit signaling, the ABCD signaling bits can be copied out of each channel before the echo canceller and

Echo Canceller Background (continued)

POTS Call Control Signaling Transport and Echo Cancelling (continued)

then reinserted after the echo canceller. The echo canceller treats the robbed bits passing through just as very low level noise since the signaling bits in each transmission direction are not necessarily related. The robbed bits are very likely to be mangled by the echo canceller process and wiped out when the nonlinear processor and comfort noise are activated. Hence, they must be reinserted before continuing in the send direction.

For older interoffice digital carrier systems connecting local switching machines within a network operator area, the situation for local calls is/was the same as for the digital loop carrier. The call control information has/had to be carried along with the voiceband payload on a call-by-call basis. For modern networks, once a call has progressed into a modern digital switch, possibly on its way to a toll network, the call control information extracted and is routed to an entirely separate signaling network. (Look up signaling system 7 in a telephone system reference.) At the receiving end, the signaling must be reinserted into the DS1/E1 carrier if a digital loop carrier is involved.

Network echo cancellers systems are often deployed at the edges of long haul networks and the local telephone networks. If, as mentioned above, all the call control signaling for the traffic that the echo canceller will see is transported in the separate signaling network, the echo canceller would not have to deal with preserving either DS1 robbed bit signaling or E1 common signaling channels.

Delay Sources: Traditional Distance, Speech Compression, Packet Switching, Error Correction, Etc.

The main source of signal delay in traditional telephone circuits came from the simple fact that nothing travels faster than light in vacuum, about 186000 miles per second, or 300000 km/s. This gives one millisecond of delay for each 186 miles or 300 km traveled. Radio waves in air travel at essentially vacuum light speed. Electrical signals in wire or coaxial cable, or light waves in optical fiber travel at slower speeds. Electrical signals in wire travel at about one-half the light speed, or about 100000 miles per hour, depending on the cable insulation and structure. This gives about 1 ms of delay for each 100 miles one way, or 2 ms for round trip delay for an echo. Older electronic terminal equipment, switches, amplifiers and the like contributed much smaller delay than the distance. A long distance call of 1500 miles would have a round trip delay of about 30 ms, and thus needs echo control.

Now, there are new sources of delay in the terminal and transmission equipment that result in echo control being needed even for short distances. Low-bit-rate voice (LBRV) compression techniques to use 8 kbits/s or even less for voice circuits require time to perform the compression and decompression. LBRV techniques tend to be used in wireless telephony where bit rates and analog bandwidth to the mobile transceivers is a precious resource.

Packet voice transmission, through ATM, the Internet or other packet networks, means that the voice signal is broken up into small packets before being transmitted. Each packet may travel a different path to the destination and experience different delays through switches and transmission links. Packets may arrive at the receiving end out of order. The packet reassembler must account for the maximum differential delay of the packets before reconstructing the signal. If a packet voice transmission system also uses LBRV, the echo control concerns may be compounded when the packet network is linked to traditional local telephony networks.

Large scale integration has made not just digital error detection feasible, but also error correction for transmission errors that occur between sender and receiver. Forward error correction requires storing long strings of input at the transmitter, and inserting redundant data codes and structures. At the receiver, the input string must all be analyzed, errors detected, and then the corrected bits substituted. The bits on the data stream may also be interleaved, or very carefully mixed before transmission. If a burst of interference produces errors in a long block of data, the deinterleaving process at the receiver distributes the errored bits as single bit errors which are easier to detect and then correct. All of this takes time and means delay in the terminal equipment. For some types of services, such as one-way television and data transmission, such delays are of much lesser concern than error-free reception. When two-way voice links are run through error-corrected links, echo control can become a concern.

Echo Canceller Background (continued)

Echoes in Telephone Calls: Electrical and Acoustic

As mentioned in other sections, the main source of electrical echoes in the telephone network is the two-to-four-wire hybrid circuit. These days, with the telephone system being essentially almost all digital once past the customer interfaces to the networks, there will be only two hybrids, one for each customer. Once the call is set up by the switching machines, there will usually be an echo canceller at each end of the long distance portion of the circuit, with each canceller taking care of the nearest hybrid echo returns. Once some signal energy appears on the channel, the cancellers can begin to train to the loss and delay characteristics of its hybrid and attached loop and customer terminals. The delay for a hybrid tends to narrow in time with only a few milliseconds where the echo occurs, say 2 ms or 3 ms within in the 64 ms capability range of most echo cancellers like the TECO3264. The characteristics of the hybrids tend to change very little during the call. Once the cancellers have trained, only minor fine tuning may be required for the rest of the call. The cancellers are monitoring to detect any changes that do occur.

Multiparty conference calls are a major and real exception to the two hybrids per call, since there will be as many hybrids as conference ports connected. One bad hybrid without echo control in a call with a dozen parties is one too many.

However, electrical echoes from hybrids are not the only sources of echoes for telephone calls. Speakerphone telephone sets are becoming more popular. The incoming sound from the phone is broadcast out into the listener's room and can bounce right back into the microphone. With the traditional handset, low level sound was fed into the ear and the head tends to block any getting to the microphone and going back out. Handsfree speakerphone/ microphone setups are also becoming more and more popular for cellular telephones in automobiles for safety reasons. Again the sound coming in can bounce around inside the car and get back to the microphone.

The positions of the speaker and microphone and the levels for speakerphones and handsfree cellular phones must be carefully done to minimize these acoustic echoes. There may be several possible reflectors in a room or car, producing multiple acoustic echoes. These acoustic echoes can vary a lot in level and delays during the call as people and reflecting objects move around. Network echo cancellers, like the TECO3264, are optimized to provide the maximum performance for the relatively stable electrical echoes. Even so, they can provide some assistance in cancelling acoustic echoes, perhaps 10 dB and more. Acoustic echo cancellers are designed to accommodate the more variable characteristics of sounds bouncing around in enclosures.

These cancellers are being included in the speakerphones and other handsfree devices. They normally have to deal with only a few ms of delay, not the 64 ms of the network cancellers. The target for attenuation of the acoustic echo is 40 dB to 45 dB between the physical acoustic path between the speaker and the microphone and the electrical echo control circuit.

Echo Canceller Background (continued)

Metallic Pair Crosstalk and Echo

Just to be complete, there is another source of voice frequency electrical echo that is usually small compared to hybrid echo. Outside of customer premises wiring, it is also becoming obsolete. However, since the four-wire interface is migrating out to the customer, it may be of more concern than in the past. There can be electrical crosstalk between the metallic pairs carrying signals in each direction for a true four-wire metallic transmission circuit. The output signal to the transmit wire pair couples back into the receive pair and adds to the received signal. This is a result of the inductive and capacitive coupling between pairs in the vicinity of one another. The level of this crosstalk echo depends on the signal frequency, the relative closeness of the pairs and the distance along the path that the pairs are close to each other. The twisting of the individual pairs, the placement of the pairs and careful wiring practices can easily keep this echo to an ERL of 40 dB and more down relative to the passing signals at voice frequencies. Echoes from hybrids and impedance mismatches in two-wire links are usually much larger.

Some customer premises, four-wire circuits do not use twisted pairs. Quadded cable is made with four nontwisted conductors in one cable sheath. Two of the conductors are used for transmission in one direction and the remaining two are used for the other direction. Without twisting, the pair-to-pair crosstalk is often 10 dB or more worse than for twisted pairs. So if a voice circuit has long runs of quad cables at the customer ends, the circuit is four-wire all the way with no hybrids, and acoustic echo cancellation is used at the ends, then the crosstalk echo component could become the main source of echo. This could be the case for some of the new telephone networks where the two-wire loop hybrid interfaces at the central offices are bypassed. If this does become a problem, an electrical echo canceller such as the TECO3264 would be a solution.

Unlike echo from a hybrid where the echo is from a point source with a very small spread in delay, this crosstalk echo coupling is distributed along the wire pairs as they run near each other. Thus, the echo is spread out in time as the signals propagate along the wires. Voice frequency signals in wire travel at about half vacuum-light-speed, or about 150 meters/µs. If the crosstalk takes place along a 3000 meter run of cable, the crosstalk is spread out for 20 µs. This would still be represented in only one of the 125 µs samples of the 64 kbits/s DS0 channels that the TECO3264 echo canceller sees. The signals experience loss as they propagate, and thus, the echo level also varies. (Hybrid echo also varies with frequency across the voiceband because of the varying level of match/balance of the compromise balance network with the connecting loop impedance. For simplified test purposes, a hybrid and the intermediate circuitry is often modeled with a flat echo return loss and a single delay total tail-end delay time, such as 6 dB ERL and 22 ms.)

Metallic wire end-to-end loss increases roughly as the square root of the frequency and crosstalk coupling tends to increase roughly as the 3/2 power of the frequency. Crosstalk is an increasing concern as the signal frequency increases. At voice frequency, crosstalk among twisted pairs is there, but the level is usually well below the echo from hybrids and impedance mismatch reflections. Thus, this crosstalk echo component should rarely be of concern for past network electrical voiceband echo canceller applications. (Nontwisted pairs, sometimes used in customer premises wiring, can result in audible crosstalk between telephone calls in premises with more than one line.) Pair-to-pair crosstalk is a major system performance limiting concern for the echo cancellers used in digital subscriber lines (DSL) where the signal frequencies reach into the MHz range.

Four-wire analog circuits in telephone networks are becoming scarce with new facility installations being four-wire digital paths. There are many left in service in the embedded plant because they still work.

Echo Canceller Background (continued)

Linear, Nonlinear, Nonlinear Processors, µ-law, A-law

The nonlinear processor is an important complement to the echo cancellation function that has been carried over from echo suppressor system design. The adjective nonlinear is used for good mathematical terminology reasons. Further, the µ-law and A-law coding and encoding rules for the payload signal carried in a 64 kbits/s DS0 signal are also nonlinear processes.

For a system with an input and an output to be called linear it must obey a rule: If the input signal is multiplied by some factor before entering the system, the output of the system will be multiplied by that same factor. A result of this property is that no frequency component will appear in the output that does not appear in the input signal as viewed in the frequency domain. Some components may appear in the output with nearly zero amplitude, i.e., they have been filtered out by the system. Even these tiny components increase or decrease with the input level. If an electronic system is not linear, the frequency components of the input signal are mixed by the system to produce new components that are the sum and difference frequencies of the input. The multiplied input yields multiplied output rule is no longer valid.

When the echo cancellation has done its best (over 30 dB of echo return loss enhancement for the TECO3264) human hearing can still detect speechlike signals coming back. Something else must be done for a final clean up. The nonlinear processor comes into play. This function is a holdover from older echo suppressor system design and can actually cut off the return path completely so nothing gets through when near-end speech is not present. When the nonlinear processor opens the switch back toward the far end on the E port, it can insert digital code word corresponding to silence. Another common term used for the nonlinear processor is center clipper since it clips out the center or the low level residual signal after the echo canceller has done its best.

Telephone customers expect to hear something even when no one is talking. So the nonlinear processor inserts comfort noise in the form of wideband white noise at about the level of the background noise as measured when no one is talking. It can be fancier and use generated noise that has some time and frequency correlation to the original background so that the transition from actual noise to artificial noise is less noticeable. When the comfort noise is inserted, the far end customer can talk loudly or softly, with a low deep bass voice or a high tenor, and the output of the nonlinear processor does not change. That is nonlinear behavior.

Electronic amplifiers are never perfectly linear, but they can be very good and acceptable for a given purpose over a wide range of input and output amplitudes and frequency ranges. A high-fidelity stereo amplifier can be good from a few Hertz up to 30 kHz and from millivolts to volts input with careful design. Telephone amplifiers do not need such a wide frequency range to transport acceptable human speech. A frequency range from 300 Hz to 3000 Hz has provided adequate speech fidelity for decades. (There are some networks that try to provide a somewhat wider frequency range, higher and lower, for a more natural sound.) A narrower range of amplitudes also eases design complexity and cost.

For very small amplitude input signals, the amplifier device turn-on/turn-off characteristics may tend to be nonlinear, or the signal may be small enough to be lost in the internal noise generated by the amplifier itself until some higher level is reached. Above some input level, the amplifier may be very linear. For some high levels of input or output, the amplifier may become nonlinear as the signal level approaches the amplifier stage supply voltage level. The range from smallest linear amplitude to the largest linear amplitude is called the amplifier's dynamic range, usually expressed in dB. The dynamic range is sometimes measured by the level range for which an dB input in the input level results in a dB increase in the output. For example, for the input level for which a one dB increase results in only a 0.5 dB increase in the output level may be a definition of the upper limit for acceptable linearity. One way to express the linearity/nonlinearity of an amplifier is by its total harmonic distortion which is a measure of the new frequency components that appear in the output and their relative magnitude to the original signal component output portions. A high-fidelity amplifier may have a harmonic distortion of a very small fraction of a percent at its maximum specified input/output level. For speech, the linearity can be a few percent distortion and still be virtually undetectable, so telephone amplifiers can economize there also. Of course, a narrow bandwidth and relatively high distortion will wreak havoc and set limits on fancy schemes to cram ever more digital data into a communications channel optimized to the hilt for speech. These high rate voiceband data techniques use very small granulations in amplitude and phase to differentiate distinct valid data points and cry for all the bandwidth possible to maximize data throughput. For human speech, the ear-brain processor seems to ignore phase information.

Echo Canceller Background (continued)

Linear, Nonlinear, Nonlinear Processors, µ-law, A-law (continued)

When the digital transport for telephony was being developed in the 1950s and early 1960s, bits were precious and higher bit rates translated to higher analog bandwidths on the copper wire transport and thus more loss. Economical, high-quality voice/speech transport was the driving goal. Digital data carried within a voice channel was at such blinding speeds as 110 baud and didn't really come into stretching what was needed for voice considerations. As mentioned, 3000 Hz is high enough for good quality speech. Shannon's sampling theorem says that one must sample at some rate greater than twice the maximum signal content frequency to fully represent a signal. A sampling rate of 8000 samples per second was a logical choice. That was easy and quick. The next question was how many bits are needed per sample. To shorten a longer story, 8 bits per sample is plenty for good speech, but not with equal-sized encoding steps per digital sample. Multiplying 8000 samples per second by 8 bits per sample yields the now familiar 64 kilobit per second DS0 voice channel rate. There was still the problem of a wide range of volume levels in human speech and that strained the eight linearily encoded bits per sample. Also, the network was entirely analog. It would be a long, long time before we had the universal digital interswitch transmission and digital switch network of today. (It is still analog from the switches to the vast majority of customers, of course.) It was realized that a long distance call that crossed the continent would be made up of several trunk links, some analog and some digital. A speech signal would have to undergo several analog-to-digital encodings and then digital-to-analog decodings to make it. The digital encoding/decoding scheme would have to be very robust and still provide toll-quality speech. The answer is the µ-law coding scheme that compressed large signals by using large sample steps and provided fine grain level resolution sample steps for very small, quiet speech. When the encoded signal was decoded back to analog speech, the linear multiplied input yields multiplied-output rule is violated, but it sounds good.

For any pulse code modulation technique for analog-to-digital conversion, there is some distortion since each coded digital word corresponds to a range of analog input level. For an original analog sample level near the bottom or top of a given level range for a given digital code, it gets decoded back to the center of the range at the range in the digital to analog conversion. The center assumes the reference levels are set exactly the same at each end, of course, but at any rate a range in becomes a point level value out.) This introduces some nonlinear distortion. The more bits in the digital code words, the smaller the end-to-end A-to-D-to-A distortion. For µ-law and A-law, the out-of-voiceband distortion products are filtered out by the sampling filters. The small inband distortion components are not noticed in the wideband speech signals.

The Europeans tweaked the µ-law coding levels just a little to develop their A-law coding scheme to provide a more constant signal-to-quantization-distortion ratio for various levels of signals. The µ-law and A-law yield a net result with only 8-bit samples that is equivalent to 11 or 12 linearily coded bits in terms of dynamic level range. This was well worth the effort to save the bits all through the switching and transmission systems. Most of the telephone networks outside the U.S.A. use A-law speech companding.

Digital echo canceller systems like the TECO3264, depend on linear PCM samples to do all of the calculations for the digital signal processing. Inside the echo canceller, the µ-law and A-law samples must be converted to linear PCM before calculations can be done. In order to preserve arithmetic accuracy, 16 linear bits are used, not just the 11 or 12 that would be enough for human hearing. (There are some music lovers with highly sensitive hearing who claim that the 16-bit linear coding used for audio compact disc [CD] recordings is not good enough. Indeed some say that only analog with infinite bit resolution is good enough. Some recent music CDs mention 20-bit processing.) For the receive side of the canceller the far-end X input encoded bits are passed to the near-end C output port. The copy of the X signal that is passed to the convolution update processor (CUP) is converted to linear PCM. For the send side of the canceller, the near-end Y input signal is converted to linear PCM. The linear PCM estimate of the echo is subtracted from the linear version of the Y signal. The difference, or residual echoes, signal is then converted to µ-law or A-law before being sent out the E port. Usually, the raw, residual echo is either attenuated further or replaced with the comfort noise when there is no double talking.

Echo Canceller Background (continued)

Adaptive Delta Modulation Pulse Code Modulation (ADPCM)

Adaptive delta modulation pulse code modulation (ADPCM) is another common standard technique for encoding voice signals with 32 kbits/s ADPCM being most common. ADPCM works by encoding not the absolute level for each time sample, but the difference between the current sample and the previous sample. With this approach, 32 kbits/s of digital bandwidth can achieve nearly the performance of 64 kbits/s A-law and μ-law. This factor of two bandwidth savings is important in some applications such as rural digital loop carrier systems and some wireless telephony systems. The TECO3264 is not designed for ADPCM since it is expected to be used primarily in the toll network. External conversion circuits between 32 kbits/s ADPCM and A-law and μ-law can be included if necessary.

Digital Speech Quality Categories

Digitally encoded audio is grouped in four general quality categories: broadcast, toll (a.k.a. network), communications, and synthetic. Broadcast usually has analog bandwidths from 5 kHz to 15 kHz depending on the service (AM radio, FM radio, television, etc.) and relatively flat frequency response requirements. Linear PCM encoded broadcast usually takes more than 64 kbits/s before any compression, such as MUSICAM for MPEG, is applied. Toll quality includes μ-law and A-law encoding at 64 kbits/s and can be achieved with an increase in distortion with 16 kbits/s ADPCM. Communications quality speech is sufficient for high reliability information transfer, but with some loss in naturalness. Companding schemes with less than 16 kbits/s and more than 4 kbits/s usually fall in this category. Digital wireless cellular systems such as GSM use 13 kbits/s speech companding. Customers complained about the 8 kbits/s companding used in the early deployments. Synthetic speech, at the current state of the art having less than 4 kbits/s bandwidth, has obvious computer or robot characteristics. See reference 6.

Quantization Distortion Units (qdu)

The distortion resulting from one ideal analog speech to 8-bit μ-law (or A-law) to analog encoding and decoding process is defined (ITU-T Recommendation G.113) as the standard against which other digital speech processing schemes are measured. The term is quantization distortion unit (qdu) for one μ-law (or A-law) encoding and decoding and corresponds to a 35 dB signal to distortion ratio. Simply inserting a digital loss pad is allocated 0.7 qdu. Thus, when the TECO3264 achieves nearly 34 dB ERLE after internal conversions, D/A conversion before the network hybrid, A/D on the return from the hybrid and finally echo cancellation; that is, all that can be done in an ideal situation.

For comparison, a 32 kbits/s ADPCM encoding and decoding is allocated 3.5 qdu. Low-bit-rate voice companding schemes have even higher distortion than 32 kbits/s ADPCM.

In general, the lower the bit rate, the higher the qdus. The planning rule for an international telephone connection is to have no more than 14 qdu for end-to-end distortion from digital processes alone, not counting such analog impairments as loss versus frequency (attenuation) distortion, background noise, etc.

High-Speed 56 kbits/s Voiceband Modems

The new 56 kbits/s high-speed voiceband modem schemes attempt to match their multiple coding levels to the discrete levels of μ-law and A-law to achieve the maximum rate possible through voice frequency designed channel toward the customer. Since the interoffice network is almost entirely digital now, only one analog-to-digital conversion and one digital-to-analog reconversion is experienced on many telephone calls, not the 2, 3, or 4 tandem encodings of the past. Voiceband modem designs are trying to work with the existing digital voice coding schemes instead of trying to fight them by using every data compression scheme ever invented. The catch is still that call control signaling (on-hook/off-hook, ringing, etc.) in North American digital carrier systems steal the least significant bit of the DS0 signal every sixth frame of the 1.544 DS1 signal to use for call control purposes. This scheme

Echo Canceller Background (continued)

High-Speed 56 kbits/s Voiceband Modems (continued)

is called robbed-bit signaling, and is an in-band control signaling transport scheme since the voice and call control share the same DS0. Thus, only 7 bits of the 8 bits per channel have a chance to get through intact, yielding a net maximum modem rate of 56 kbits/s, not 64 kbits/s. Further, the U.S. Federal Communications Commission (FCC) signal power limits result in an actual maximum rate of only 53 kbits/s. The 56 kbits/s voiceband modems can tolerate only one analog to µ-law conversion. The new V.90 standard for 56 kbits/s modems is nearing final approval.

The 56 kbits/s data transport is asymmetric; that is, the 56 kbits/s maximum rate is only available from the service provider to the customer. The customer to service provider direction uses normal voiceband modem techniques, such as V.34.

E1 based transmission systems do not use robbed-bit inband signaling, but carry all of the call control for 31 voice circuits in one of the 32 DS0 slots for a common control channel, usually channel 0. This is called out-of-band control signaling transport.

Companding

The term companding is made up from two words, compressing and expanding. Companding in telephony is a technique used to preserve the dynamic level range of a signal such as speech when it must be transmitted through an electronic channel that has less dynamic range and may distort low levels or high levels. The compressor circuit at the transmitter end reduces the input level range by a stated rule. The expandor circuit at the receiver end restores the original dynamic range by applying the inverse of the compressing rule. A-law and µ-law compandors work by allocating more bits for low-level signals than for high-level signals, thus preserving accuracy and a low background noise at the cost of increased nonlinearity for loud signals for which human hearing is more tolerant. Other speech companding schemes include those which work on a syllable-by-syllable basis, and those that use input/output ratio processing like 1 dB output increase for 2 dB input increase (and reverse at the receiver). Low-bit-rate voice algorithms employ various level companding techniques to help reduce the data rate.

Companding can be divided into lossy or lossless categories in terms of how faithfully they can restore the original information at the receiving end. Lossy companding is often used for speech, music, and video data compression, and lossless for digital data archival. Lossless companding is used for digital data files, such as source code, numerical information, and executable code, where every bit is important. Lossless compandors can shrink the size of computer files and then restore them with every bit intact. zip, arj, tar, and lzh are some of the lossless compandor techniques used for computer data archival and transport.

Lossy companding techniques discard as much information as possible and still reconstruct a signal at the end that is good enough for the given purpose for human ear or eye and brain processing. Specific lossy techniques include: signal frequencies limited at low and high ends (linear), continuous level ranges reduced to discrete stairsteps (nonlinear), low-level signal components near in frequency to high level components discarded by digital filtering (linear). Some lossy companding schemes include: µ-law and A-law for telephony, MUSICAM for music audio storage and transmission, MPEG for video, JPEG and GIF for still images.

References

1. *Concentration Highway Interface (CHI)* Interface Specification, document number DS90-124SMOS, Lucent Technologies Microelectronics Group.
2. *Pulse Code Modulation (PCM) of Voice Frequencies*, ITU-T Recommendation G.711.
3. *Echo Suppressors*, ITU-T Recommendation G.164.
4. *Echo Cancellers*, ITU-T Recommendation G.165.
5. *Digital Network Echo Cancellers*, ITU-T Recommendation G.168.
6. *Speech Coding: A Tutorial Review*, A. S. Spanias, Proceedings of the IEEE, Vol. 82, No. 10, October 1994, pages 1541 to 1582.
7. *Telecommunications Transmission Engineering*, Volumes 1—3, Bellcore, ISBN 1-97808-04-02, 1990.
8. *Packet-Based Multimedia Communications Systems*, ITU-T Recommendation H.323, Geneva, 1998.

For additional information, contact your Microelectronics Group Account Manager or the following:
INTERNET: **http://www.lucent.com/micro**
E-MAIL: **docmaster@micro.lucent.com**
N. AMERICA: Microelectronics Group, Lucent Technologies Inc., 555 Union Boulevard, Room 30L-15P-BA, Allentown, PA 18103
1-800-372-2447, FAX 610-712-4106 (In CANADA: **1-800-553-2448**, FAX 610-712-4106)
ASIA PACIFIC: Microelectronics Group, Lucent Technologies Singapore Pte. Ltd., 77 Science Park Drive, #03-18 Cintech III, Singapore 118256
Tel. (65) 778 8833, FAX (65) 777 7495
CHINA: Microelectronics Group, Lucent Technologies (China) Co., Ltd., A-F2, 23/F, Zao Fong Universe Building, 1800 Zhong Shan Xi Road, Shanghai 200233 P. R. China **Tel. (86) 21 6440 0468**, **ext. 316**, FAX (86) 21 6440 0652
JAPAN: Microelectronics Group, Lucent Technologies Japan Ltd., 7-18, Higashi-Gotanda 2-chome, Shinagawa-ku, Tokyo 141, Japan
Tel. (81) 3 5421 1600, FAX (81) 3 5421 1700
EUROPE: Data Requests: MICROELECTRONICS GROUP DATALINE: **Tel. (44) 7000 582 368**, FAX (44) 1189 328 148
Technical Inquiries: GERMANY: **(49) 89 95086 0** (Munich), UNITED KINGDOM: **(44) 1344 865 900** (Ascot),
FRANCE: **(33) 1 40 83 68 00** (Paris), SWEDEN: **(46) 8 594 607 00** (Stockholm), FINLAND: **(358) 9 4354 2800** (Helsinki),
ITALY: **(39) 02 6608131** (Milan), SPAIN: **(34) 1 807 1441** (Madrid)

microelectronics group

Lucent Technologies
Bell Labs Innovations

August 1999
MN98-071TIC